Amendments to the Drawings

The attached sheet of drawings includes changes to Fig. 1I. This sheet, which includes Figs. 1I and 2, replaces the original sheet including Figs. 1I and 2. In Fig. 1I, the SEQ ID NO. has been inserted to identify the corresponding listed sequence.

Attachment: Replacement Sheet
Annotated Sheet Showing Changes

Remarks

Applicants respond to the Office Action dated August 18, 2008, for which a three month period of response is given. Upon entry of the foregoing amendment, claims 1, 3-5, 7-8, 16-19 and 34 will be pending in the application. Claims 2, 6, 9-15, 20-33 and 35-40 have been canceled.

The specification has been amended to insert a reference to SEQ ID NO: 45-48 after the corresponding amino acid sequences throughout the specification. Support for the amendment is found on page 46, lines 21-25. Support for the amendment on page 109 can be found in the specification as originally filed.

In amended Fig. 1I, the previously omitted SEQ ID NO. 2-35 have been added. Support for the amendment can be found at page 51, line 14 through page 52, line 15 of the specification.

Sequence Rules

Applicants respectfully submit that in the amended claims, specification and drawings, the amino acid sequences are indentified by SEQ ID Nos. All of the sequences appearing in the claims, specification and drawings are included in the Sequence Listing as originally filed on December 12, 2005.

Claim Objections

Claims 1-19 and 34 have been objected as reciting non-elected subject matter. In view of the amendments to the claims, Applicants respectfully submit that the objection has been overcome.

Claim Rejections – 35 USC §112

Claims 3 and 4 have been rejected under 35 U.S.C. §112, first paragraph, as being noncompliant with the enablement and written description requirements. The Examiner contends that the specification does not enable a person skilled in the art to make the invention where the natural α-glucan phosphorylase has at least 50% identity with position 1 to position 916 of SEQ ID NO: 2 (claim 3) or where the amino acid sequence of the natural α-glucan phosphorylase is encoded with a nucleic acid

molecule which hybridizes under stringent conditions (claim 4).

Regarding claim 3, the Examiner states that the specification does not support the broad scope of the claims because the specification does not establish (A) regions of the protein structure which may be modified without effecting α-glucan phosphorylase activity, (B) the general tolerance of α-glucan phosphorylase enzyme to modification and extent of such tolerance, (C) a rational and predictable scheme for modifying any α-glucan phosphorylase enzyme residues with an expectation of obtaining the desired enzymatic or biological function and being thermostable, and (D) which of the essentially infinite possible choices is likely to be successful. Regarding claim 4, the Examiner contends that the specification does not sufficiently define the stringent conditions under which the hybridizations are to take place and that given the unpredictability of the art and the nature of hybridization experiments in general, it is not sufficient to merely cite hybridization without clear and explicit recitation of the conditions associated with the hybridization.

With regard to the written description requirement, the Examiner contends that there is no information in the disclosure such as drawings or structural formulas that would indicate which amino acids can be varied from SEQ ID NO: 2 in the claimed genus and still retain the catalytic activity. The Examiner contends that this lack of disclosure along with the lack of knowledge and predictability in the art would not indicate to a person skilled in the art that applicant was in possession of the claimed genus based on the disclosure of several naturally occurring proteins having α-glucan phosphorylase enzyme activities without guidance to specific modifications.

Applicants respectfully disagree with the Examiner's contentions. The present invention relates to an α-glucan phosphorylase having improved thermostability, which is obtained by modifying a natural α-glucan phosphorylase. Claim 3 recites a natural α-glucan phosphorylase which is used for modification. As described on page 29, lines 26-29, "α-glucan phosphorylase" means an enzyme having α-glucan phosphorylase activity. Thus, the natural α-glucan phosphorylase recited in Claim 3 has α-glucan phosphorylase activity and at least 50% identity with an amino acid sequence of position 1 to position 916 of SEQ ID NO: 2. Claim 3 is not intended to cover a non-natural α-glucan phosphorylase being modified based on the amino acid sequence of

SEQ ID NO: 2 in the natural α-glucan phosphorylase.

As described on page 31, lines 6-8, it is thought that α-glucan phosphorylase is ubiquitously present in various plants, animals, and bacteria which can store starch or glycogen, and it is thought that the plants, which can store starch or glycogen, have an active α-glucan phosphorylase.

Regarding the naturally occurring α-glucan phosphorylase, Applicants submit the attached Exhibit 1: Structure and properties of *Thermus aquaticus* α-glucan phosphorylase expressed in *Escherichia coli*. T. Takaha, M. Yanase, H. Takata and S. Okada. *J. Appl. Glycosci.*, (2001) 48, 71-78. Exhibit 1 describes the alignment score (%) between the amino acid sequences of glucan phosphorylase from various sources (e.g., bacterial, plant, and animals). Exhibit 1 also describes that glucan phosphorylase obtained from potato has a very low similarity (e.g., 9 and 10 %) to glucan phosphorylase obtained from bacteria (see Fig. 5 on page 77 of
Exhibit 1). Moreover, the amino acid sequences of glucan phosphorylase derived from similar species having similar properties have a relatively high similarity of 30-40%.

As described in Table 4 on page 66 of the present specification, the percent identity (%) between the amino acid sequences of α-glucan phosphorylase derived from plants are very high and are 57% or more with regard to 15 α-glucan phosphorylase shown in Table 4. Therefore, it would have been known to one of ordinary skill in the art, based on an active natural amino acid sequence, to obtain an active modified amino acid sequence by modifying one or several amino acid acids. Furthermore, the amino acid residue essential for glucan phosphorylase activity on the amino acid sequence of glucan phosphorylase was known in the art at the time of filing the present application. Applicants submit the attached Exhibit 2: Evolution of allosteric control in glycogen phosphorylase. John W Hudson, G. Brian Golding and Michael M Crerar. *J. Mol. Biol.*, (1993) 234, 700-721.

Exhibit 2 indicates the comparison of many amino acid sequences of glucan phosphorylases (also known as Glycogen Phosphorylase). At the highlighted portion in Fig.1 on page 707, the motif sequence 3L "RIVKFITDV" is indicated in the sequence of PotL. Fig. 1 indicates that amino acid residues at the site of motif sequence 3L are not involved in glucan phosphorylase activities and are not essential for glucan

phosphorylase activities. Furthermore, Exhibit 2 indicates the amino acid residue which is essential for glucan phosphorylase activity.

Thus, the α-glucan phosphorylase having improved thermostability recited in claims 1 and 3 is fully supported in the originally filed application.

With regard to claim 4, Applicants have amended claim 4 to recite the details of the stringent conditions. The specification states that the selection of appropriate stringent conditions is well-known to those skilled in the art (page 67, lines 4-22) and describes a specific example of suitable stringent conditions (see lines 10-22). Thus the disclosure provides sufficient guidance to enable a person skilled in the art to make the invention of claim 4 in a manner reasonably correlated with the scope of the claims without undue experimentation. Applicants respectfully request withdrawal of the rejection of claims 3 and 4 under 35 U.S.C. §112, first paragraph.

Claims 1-19 and 34 have been rejected under 35 U.S.C. §112, second paragraph, as being indefinite. The Examiner contends that the phrase "derived from a plant" or "plant derived" are vague in the context used. As suggested by the Examiner, Applicant has amended claims 1, 8 and 34 to recite the phrase "obtained from a plant". Applicants respectfully request withdrawal of the rejection of claims 1, 3-5, 7-8, 16-19 and 34 under 35 U.S.C. §112, second paragraph.

Claim Rejection – 35 USC §101

Claims 1-19 and 34 have been rejected under 35 U.S.C. § 101 as being directed to non-statutory subject matter. The Examiner states that "in the absence of the hand of man," naturally occurring proteins and/or nucleic acids are considered non-statutory subject matter, but that the rejection may be overcome by amending claims 1 and 34 to recite wording such as "an isolated α-glucan phosphorylase".

Applicants respectfully disagree, as the current claim language does not recite a naturally occurring protein or nucleic acid but rather recites a protein that is obtained by modifying a natural protein. The claims indicate that the modified protein is obtained by modifying an amino acid residue in a particular motif sequence. Thus, the claims recite a protein that has been modified by the hand of man. Applicants respectfully request

withdrawal of the rejection under 35 U.S. C. §101.

Conclusion

In view of the foregoing amendments and remarks, it is believed that the application is in condition for allowance and a notice of allowance is therefore respectfully requested.

In the event any fee or additional fee is due in connection with the filing of this paper, the Commissioner is authorized to charge those fees to our Deposit Account No. 18-0988 (under Docket Number YAMAP0997US). In the event an extension of time is needed to make the filing of this paper timely and no separate petition is attached, please consider this a petition for the requisite extension and charge the fee to our Deposit Account No. 18-0988 (under Docket Number YAMAP0997US).

Respectfully submitted,

RENNER, OTTO, BOISSELLE & SKLAR, LLP

By /Heidi A. Boehlefeld/
Heidi A. Boehlefeld, Reg. No. 34,296

1621 Euclid Avenue
Nineteenth Floor
Cleveland, Ohio 44115
(216) 621-1113

Exhibit 1

[*J. Appl. Glycosci.*, Vol. 48, No. 1, p. 71–78 (2001)]

Structure and Properties of *Thermus aquaticus* α-Glucan Phosphorylase Expressed in *Escherichia coli*

Takeshi Takaha,* Michiyo Yanase, Hiroki Takata and Shigetaka Okada

Biochemical Research Laboratory, Ezaki Glico Co., Ltd.
(4–6–5, Utajima, Nishiyodogawa-ku, Osaka 555–8502, Japan)

The α-glucan phosphorylase gene from *Thermus aquaticus* was isolated using partial amino acid sequences of purified enzyme. The identity of the gene was confirmed by expression in *Escherichia coli* resulting in thermostable glucan phosphorylase activity. The open reading frame of this gene consisted of 2460 bp and encoded a polypeptide of 819 amino acids. The deduced amino acid sequence exhibits high identity (32–43%) to 7 putative and 2 characterized glucan phosphorylases, but showed weak similarity to other well characterized glucan phosphorylases from various sources. Due to its high expression level and thermal stability, the recombinant enzyme was easily purified from *E. coli* cell extracts, and employed to characterize its activity. The smallest primer molecule for a synthetic reaction was maltotriose and the smallest effective substrate for a degradation reaction was maltotetraose. These results suggest that *T. aquaticus* glucan phosphorylase, and at least 9 other enzymes, form a new group of glucan phosphorylases whose structure and substrate specificity differ from the traditional glucan phosphorylases. The purified enzyme was also employed to investigate the effect of temperature and pH on the activities in both directions. The activity exhibited a pH optimum of 8.0 for phosphorolytic reaction but 7.0 for synthesis reaction. The optimum temperature for phosphorolytic reaction was 80–85°C, while the one for synthesis reaction was 50°C.

α-Glucan phosphorylase (EC 2.4.1.1) catalyzes the reversible phosphorolysis of α-1,4-glucan and is widely distributed in microorganisms, plants and animals. All known α-glucan phosphorylases require pyridoxal 5′-phosphate for activity and seem to share a similar catalytic mechanism.[1] Although enzymes from distinct origin differ in their substrate preference and their mode of regulation, all phosphorylases belong to a large group of highly homologous phosphorylases, comprising glycogen phosphorylases from bacteria, yeast and animals, starch phosphorylases from plants, and maltooligosaccharide phosphorylases of bacteria.[2,3]

It has been reported that the smallest primer molecule for glucan synthetic reaction of glucan phosphorylase is maltotetraose, and the smallest effective substrate for glucan degradation reaction was maltopentaose, and generally believed that this is the common feature for glucan phosphorylases. However, the enzymes from *Thermus thermophilus*[4] and *Thermococcus litoralis*[5] were recently reported to have distinct substrate specificity, where maltotriose is the smallest primer for glucan synthesis and maltotetraose is the smallest substrate for glucan degradation.

An increasing number of glucan phosphorylase or hypothetical glucan phosphorylase genes are now available largely because of the rapid progress in microbial genome projects. We have found that some recently identified glucan phosphorylases showed very weak similarity to traditional glucan phosphorylases, and might constitute a new subgroup. More interestingly, the glucan phosphorylase from *T. litoralis*, which has a distinct substrate specificity as mentioned above, also belongs to this new sub-group. This may suggest the correlation between the new structure and new substrate specificity. Unfortunately, none of other enzymes

*To whom correspondence should be addressed.

belonging to this new sub-group was characterized for its substrate specificity, so the structure of glucan phosphorylase from *Thermus* spices, is of great interest to investigate this hypothesis.

The aim of this work is primarily to isolate the gene for *Thermus aquaticus*, and investigate the similarity of its primary structure with the characterized and putative phosphorylases especially with the one from *T. litoralis*. This work also aims to characterize the substrate specificity and other properties of the recombinant enzyme purified from *E. coli* in order to seek the exploitation of this enzyme for the synthesis of amylose[5] and glucose 1-phosphate.[6]

MATERIALS AND METHODS

Materials. Maltooligosaccharides were purchased from Hayashibara Biochemical Laboratories Inc. (Okayama, Japan). Unless otherwise specified, all chemicals were purchased from Wako Pure Chemical Industries Ltd. (Osaka, Japan).

Partial purification of glucan phosphorylase from T. aquaticus. *Thermus aquaticus* ATCC 33923 was grown at 70°C for 18 h in a medium containing 0.4% (w/v) yeast extract, 0.8% (w/v) tryptone, 0.2% (w/v) NaCl and 1% (w/v) glucose, pH 7.0. The cells were harvested by centrifugation and washed with 50 mM Tris-HCl (pH 7.0) containing 5 mM EDTA and 1 mM 2-mercaptoethanol (buffer A). The cells were disrupted by sonication in buffer A at 4°C, and centrifuged to remove cell debris. This crude extract was heated at 60°C for 30 min and centrifuged. The supernatant was filtered through a 0.45 μm membrane and then loaded onto a Q-Sepharose fast flow column (26×150 mm; Pharmacia Biotech Inc.) and washed with 200 mM NaCl in buffer A, then eluted with buffer A containing 300 mM NaCl. Concentrated Tris buffer, ammonium sulfate, EDTA, and 2-mercaptoethanol were added to the collected solution to give final concentrations of 50 mM Tris-HCl (pH 7.0), 500 mM ammonium sulfate, 5 mM EDTA and 1 mM 2-mercaptoethanol. The solution was loaded onto a phenyl-TOYOPEARL 650M column (16×100 mm TOSOH, Tokyo, Japan) and washed with 100 mM ammonium sulfate in buffer A, then eluted by buffer A. Active fraction was loaded onto a Source15Q column (10×100 mm; Pharmacia Biotech Inc.), and eluted with a linear gradient of 100–400 mM NaCl in buffer A. Active fraction were concentrated and applied to a Sephacryl S-300 column and eluted with 100 mM NaCl in buffer A.

Determination of amino acid sequence of glucan phosphorylase. To obtain a pure but inactivated polypeptide, partially purified enzyme was applied to reverse phase HPLC using a C4 column (250×4.6 mm; YMC Biochemicals, Japan) and eluted with a linear 0–60% (v/v) gradient of acetonitrile with 0.1% (v/v) trifluoroacetic acid, and the enzyme was collected, concentrated *in vacuo* and subjected to a peptide sequencer. Digestion of purified enzyme with trypsin and the purification of tryptic fragments were described previously.[7]

Expression of glucan phosphorylase gene in Escherichia coli. An expression plasmid pKK 388-GP was constructed as follows. In order to produce new restriction sites in 5′ and 3′ non-translated regions, the glucan phosphorylase gene was amplified with two oligonucleotide primers, 5′-AAATCGATAGGAAAACATATGAACGTCCT CGGACGGACGGATCAC-3′ and 5′-TTCTAGA CGCCCCTAGGCCCAGCGCAC-3′, using the plasmid pTFGP01 as a template. The amplified fragment was digested with *Cla*I and *Xba*I, then introduced into the *Cla*I-*Xba*I sites of pKK388-1 (Clontech) to produce pKK388-GP. *E. coli* MC 1061 carrying pKK388-GP was grown in LB medium containing ampicillin (100 μg/mL). At the late log phase, IPTG (0.1 mM) and pyridoxine (0.1 mM) were added and the cells were cultured for an additional 16 h at 37°C, then harvested by centrifugation. Further purification steps were the same as described above, but gel-filtration chromatography with Sephacryl S-300 column was omitted.

Glucan phosphorylase activity assay. Enzyme activity was determined for synthesis (assay A) or phosphorolysis (assay B).

Assay A. The production of inorganic phosphate from glucan and glucose 1-phosphate was measured by the method described by Saheki *et al.*[8] with minor modification. A reaction mixture

(200 μL) containing 100 mM Tris-HCl (pH 7.0), 1% (w/v) soluble starch, and 45 mM glucose-1-phosphate (Boehringer Mannheim, Germany) was incubated at 50°C for 30 min. The reaction was stopped by the addition of 10 μL of 20% (w/v) sodium dodecyl sulfate solution. Then 800 μL of molybdate reagent (15 mM ammonium molybdate, 100 mM zinc acetate, pH 5.0) and 200 μL of ascorbic acid reagent (10% (w/v), pH 5.0) were added to the mixture. This mixture was incubated at 30°C for 15 min, and the absorbance was measured at 850 nm. Enzyme mixed with sodium dodecyl sulfate was used as a blank. One unit was defined as the amount of enzyme that produced 1 μmol of phosphate in 1 min.

Assay B. Activity was assayed by determining the amount of glucose-1-phosphate produced from glucan and inorganic phosphate. The reaction mixture (60 μL) containing 100 mM sodium phosphate buffer (pH 7.0), 1% (w/v) soluble starch and enzyme was incubated at 70°C for 10 min. The reaction was terminated by diluting the mixture with 740 μL of 50 mM Tris-HCl (pH 7.0) which had been kept on ice. The mixture was then incubated with 400 μL of assay reagent containing 200 mM Tris-HCl (pH 7.5), 3 mM NAD^+, 15 mM $MgCl_2$, 3 mM EDTA, 15 μM glucose 1,6-bisphosphate, 6 μg /mL phosphoglucomutase (from rabbit muscle; Boehringer Mannheim) and 6 μg/mL glucose-6-phosphate dehydorogenase (from *Leuconostoc* sp.; Boehringer Mannheim). The mixture was incubated at 30°C for 30 min, and the absobance at 340 nm was measured. One unit was defined as the amount of enzyme that produced 1 μmol of glucose 1-phosphate in 1 min.

Thin Layer Chromatography (TLC). TLC of oligosaccharides was carried out according to the method described by Takata *et al.*[9] with minor modification. Reaction mixtures (20 μL) were loaded twice onto a 100 μL column of anion exchange resin (Diaion SA11AS Cl-form, Mitsubishi Kasei Co., Tokyo) to remove inorganic phosphate and glucose 1-phosphate. Three μL of the eluent was spotted on a TLC plate (Silica gel 60; Merck, Darmstadt, Germany) and developed three times in acetonitorile/water (75: 25 v/v), and detected by heating at 130°C for 5 min after spraying 50% (v/v) sulfuric acid in methanol.

RESULTS AND DISCUSSION

Purification and determination of partial amino acid sequence of glucan phosphorylase from T. aquaticus.

A cell extract of *T. aquaticus* was subjected to four chromatography steps as described in MATERIALS AND METHODS. However, the sample after Sephacryl S-300 gel-filtration chromatography still produced two bands (major and minor bands as shown in Fig. 1A) when analyzed by SDS-PAGE. In order to identify the band for glucan phosphorylase, the same sample was subjected to native-PAGE and the gel was stained for phosphorylase activity. As shown in Fig. 1B, a phosphorylase activity was only found in the major band. The major band with an estimated molecular mass of 91 kDa, was further purified by reverse phase HPLC and subjected to peptide sequencing, and N-terminal amino acid sequence (MNVLGR) was determined. To determine the internal amino

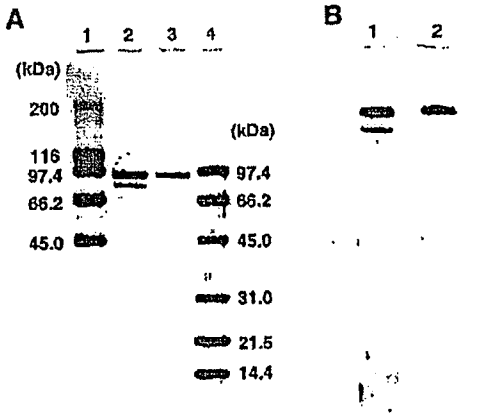


Fig. 1. SDS-PAGE (A) and native-PAGE (B) of glucan phosphorylase from *T. aquaticus*.

(A) Lane 1 and 4, molecular mass markers; lane 2, sample after Sephacryl S-300 chromatography; lane 3, sample after reverse phase HPLC with C4 column. (B) Lane 1, sample after Sephacryl S-300 chromatography stained with coomassie blue; lane 2, same as lane 1, but stained with iodine solution after incubating the gel in 20 mM glucose-1-phosphate with 0.05% (w/v) of soluble starch for 3 h at 70 °C.

acid sequence of the enzyme, the purified enzyme was digested with trypsin and four tryptic fragments were purified by reverse phase HPLC, after which additional amino acid sequences (ELVAEGYPFR, YLGGFW, HYAEVFGPEW and LALLHPVTGR) were obtained.

Isolation of glucan phosphorylase gene of T. aquaticus and its expression in E. coli.

Oligonucleotide primers corresponding to these five amino acid sequences were synthesized and used to amplify a DNA segment for phosphorylase from *T. aquaticus* genome. Each amplified DNA fragment obtained from such PCR experiment was sequenced and searched for its homology to the glucan phosphorylases genes in the database. A 420 bp PCR fragment amplified with two primers, 5′-TAYYTNGGNGGNTTYTGG-3′ and 5′-CRAA NACYTCNGCRTART-3′, showed significant similarity to glucan phosphorylase, and thus was used as a DNA probe to screen the *T. aquaticus* genomic DNA library.[10] Five positive plaques were obtained but one clone having the longest DNA insert (3.9 kb, pTFGP01) was selected for further analysis. The nucleotide sequence (DDBJ accession number: AB047267) reveals that the insert DNA (3891 bp) contained the open reading frame of 2460 bp (819 amino acids). The amino-terminal and internal amino acid sequences obtained from peptide sequencing of purified protein were all found in the deduced amino acid sequence.

Expression of this gene in *E. coli* was sought to confirm its identity. The plasmid vector pKK388-1 containing a *tac* promoter was employed to direct expression of the complete open reading frame (see materials and methods). When *E. coli* MC 1061 carrying expression plasmid pKK388-GP was grown with the inducer IPTG, thermostable glucan phosphorylase activity was detected in the soluble fraction of cell extract. This result clearly demonstrates that this open reading frame codes for glucan phosphorylase.

Purification and characterization of recombinant glucan phosphorylase.

The glucan phosphorylase was purified to homogeneity from the recombinant *E. coli* cell extracts by the purification steps summarized in Table 1. The purified enzyme produced a clear single band on SDS-PAGE with an apparent molecular mass of 91 kDa, which showed a good agreement with the enzyme prepared from *T. aquaticus* (data not shown). The purified enzyme was next employed to investigate the effect of pH and temperature on its activity and stability (Figs. 2 and 3). The optimum temperature and pH for the glucan degradation reaction were 80–85°C, and 8.0, respectively, which agreed well with the reported value for the enzyme from *T. thermophilus*. However, the optimum temperature and pH for glucan synthetic reaction were 50°C and 7.0, respectively. This drastic shift in optimum temperature and pH are important upon the exploitation of this enzyme for amylose and glucose 1-phosphate production. Similar results were also obtained in the glucan phosphorylase from *Chlorella*,[11] but not tested for the enzyme from *T. thermophilus*. The enzyme was stable after incubation at 60°C for 30 min, and about 80% of enzyme activity was retained even after incubation at 80°C for 30 min. The enzyme was stable at pH 5.5 to 9.5 at 50°C

Table 1. Purification of *T. aquaticus* glucan phosphorylase expressed in *E. coli*.

Step	Total activity[a] (U)	Total protein (mg)	Specific activity (U/mg)	Activity recovered (%)
Crude extract	3121	739	4.22	100
Heat treatment (60°C, 30 min)	2030	245	8.29	65.0
Q-Sepharose	1870	168	11.2	59.9
Phenyl TOYOPEARL	1518	117	13.1	48.6
Source 15Q	906	75.9	13.7	29.0

[a]Activity was determined in the glucan synthetic direction (assay A).

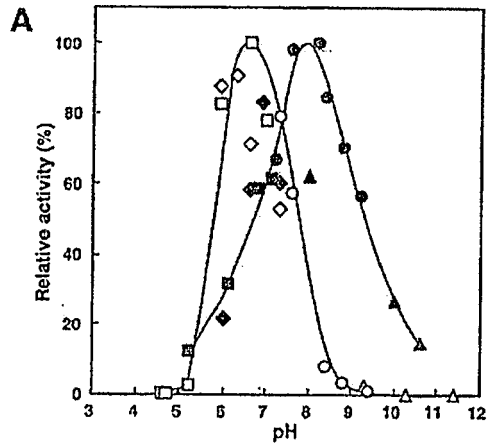


Fig. 2. Effect of pH on the activity (A) and stability (B) of *T. aquaticus* glucan phosphorylase.

(A) The activity was assayed in 0.2 M buffer at 50°C, for 30 min. (B) Purified enzyme was incubated in 50 mM buffers at 50°C for 30 min. After adjusting the pH to 7 by diluting the enzyme with 0.2 M Tris-HCl buffer (pH 7.0), the remaining activity was assayed, in synthetic (open symbols) and phosphorolytic (closed symbols) directions at 50 °C and 70°C, respectively. Buffer system used are sodium acetate (□, ■), sodium citrate (◇, ◆), Tris-HCl (○, ●) and Glycin-HCl (△, ▲). The highest activity was designated as 100%.

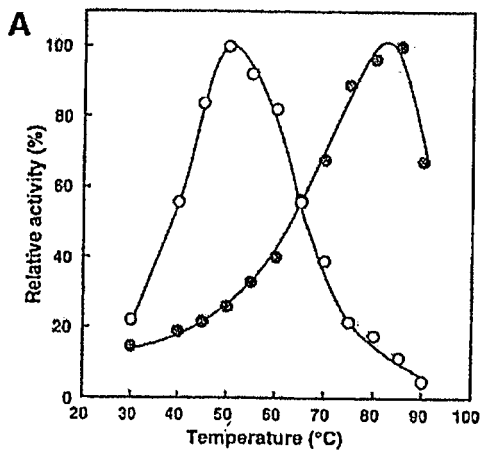


Fig. 3. Effect of temperature on the activity (A) and stability (B) of *T. aquaticus* glucan phosphorylase.

(A) The activity was assayed in synthetic (○) and phosphorolytic (●) directions at indicated temperature. (B) Purified enzyme in 20 mM Tris-HCl buffer (pH 7.0) was incubated at indicated temperature for 30 min. After the mixture was cooled, the remaining activity was assayed in synthetic (○) and phosphorolytic (●) directions at 50°C and 70°C, respectively. The highest activity was designated as 100%.

for 30 min. All these results suggest that glucan phosphorylase from *T. aquaticus* and *T. thermophilus* resemble each other in their properties.

It has been reported that the smallest primer molecule for glucan synthetic reaction of glucan phosphorylase is maltotetraose, and the smallest effective substrate for glucan degradation reaction was maltopentaose, and generally believed that this is the common feature of glucan phosphorylases. However, the enzymes from *T. thermophilus* were recently reported to have distinct substrate specificity, where maltotriose is the smallest primer for glucan synthesis and maltotetraose is the smallest substrate for glucan degradation.[4] Subsequently, a

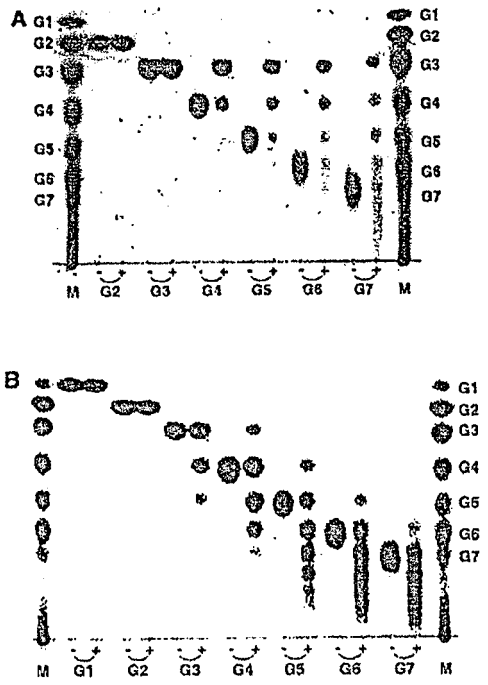


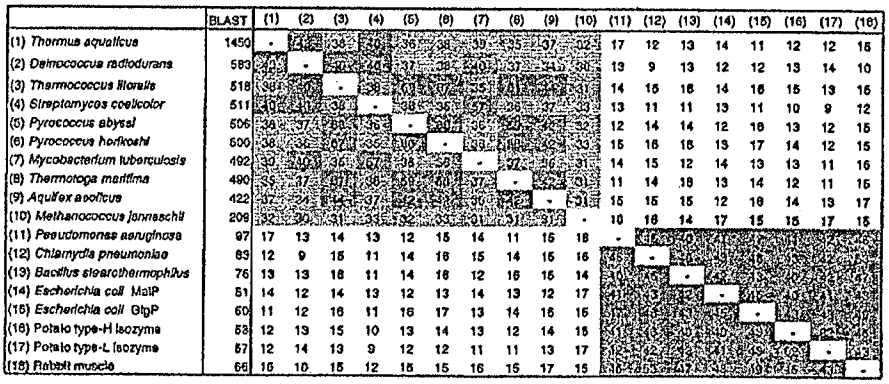
Fig. 4. TLC analyses of the degradation (A) and synthetic reaction (B) of *T. aquaticus* phosphorylase on various oligosaccharides.

(A) A reaction mixture (20 μL) consisting of 1% (w/v) oligosaccharide was incubated with (+) or without (−) enzyme (0.005 U) at 50°C for 3 h. After removing inorganic phosphate and glucose 1-phosphate using anion exchange resin, 3 μL of mixture was spotted on TLC plate. (B) A reaction mixture (20 μL) consisting of 1% (w/v) oligosaccharide, and 10 mM glucose 1-phosphate was incubated with (+) or without (−) enzyme (0.005 U) at 50°C for 3 h. After removing inorganic phosphate and glucose 1-phosphate using anion exchange resin, 3 μL of mixture was spotted on TLC plate. M, maltooligosaccharide makers; G1, G2, G3, G4, G5, G6 and G7 are glucose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose and maltoheptaose, respectively.

similar property was also reported in the glucan phosphorylase from *T. litoralis*.[2] Thus the substrate specificity of the purified enzyme was next investigated. In the presence of inorganic phosphate, the enzyme can attack maltotetraose and larger, while maltose and maltotriose were not affected (Fig. 4A). In the presence of glucose 1-phosphate, maltotriose and larger were effective primers for glucan synthesis, but glucose and maltose were not used as primers (Fig. 4B). These results indicate that the minimum primer molecule for synthetic reaction and minimum substrate for degradation reaction were maltotriose and maltotetraose, respectively, which agree with those for the enzyme from *T. thermophilus*[4] and *T. litoralis*,[2] but one unit smaller than the reported properties for traditional glucan phosphorylases from various sources (e.g. *E. coli*, potato or rabbit enzymes).

Comparison of primary structure of T. aquaticus glucan phosphorylase with other glucan phosphorylases.

Since the enzyme has distinct substrate specificity as described above, comparison of primary structure with other glucan phosphorylase is of great interest, especially to understand the structure function correlation. The deduced amino acid sequence of *T. aquaticus* glucan phosphorylase was subjected to BLAST homology search (http://www.ncbi.nlm.nih.gov/BLAST/). The eleven sequences having from the highest to the eleventh highest BLAST score were listed in Fig. 5 (numbers 2 to 12). This result suggests that the enzyme exhibits high similarity to 9 putative and characterized glucan phosphorylases from thermophilic bacteria (*Deinococcus radiodurans, Thermotoga maritima* and *Aquifex aeolicus*), mesophilic bacteria (*Streptomyces coelicolor, Mycobacterium tuberculosis*) and hyperthermophilic archaea (*Thermococcus litoralis, Pyrococcus abyssi, Pyrococcus horikoshii, Methanococcus jannaschii*) (Fig. 5, numbers 2 to 10), but shows weak similarity to other glucan phosphorylases (Fig. 5, numbers 11 and 12). The enzyme also shows low BLAST score against well characterized glucan phosphorylases from *Bacillus stearothermophilus, E. coli*, potato and rabbit muscle (Fig. 5, numbers 13 and 18). The amino acid sequences of eighteen glucan phosphorylases described above were next aligned using the CLUSTAL W (1.81) program (http://www2.ebi.ac.uk/clustalw/), and the alignment scores (%) were summarized in Fig. 5. The result clearly indicates that *T. aquaticus* enzyme and 9 other glucan phosphorylases (Fig. 5, numbers 2 to 10) form a new sub-group, since all these enzymes show higher



	BLAST	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) *Thermus aquaticus*	1450	-	44	36	40	36	34	39	35	37	22	17	12	13	14	11	12	12	15
(2) *Deinococcus radiodurans*	583	44	-	40	44	37	36	40	37	41	21	13	9	13	12	12	13	14	10
(3) *Thermococcus litoralis*	518	36	40	-	38	41	37	35	41	41	21	14	15	16	14	16	15	13	16
(4) *Streptomyces coelicolor*	511	40	44	38	-	36	35	37	39	37	23	13	11	11	13	11	10	9	12
(5) *Pyrococcus abyssi*	506	36	37	41	36	-	61	36	38	42	22	12	14	14	12	16	13	12	15
(6) *Pyrococcus horikoshi*	500	34	36	37	35	61	-	39	40	42	23	15	16	15	13	17	14	12	15
(7) *Mycobacterium tuberculosis*	492	39	40	35	37	36	39	-	37	36	21	14	15	12	14	13	13	11	16
(8) *Thermotoga maritima*	490	35	37	41	39	38	40	37	-	43	20	11	14	16	13	14	12	11	15
(9) *Aquifex aeolicus*	422	37	41	41	37	42	42	36	43	-	21	15	15	15	12	16	14	13	17
(10) *Methanococcus jannaschii*	209	22	21	21	23	22	23	21	20	21	-	16	16	14	17	15	15	17	15
(11) *Pseudomonas aeruginosa*	97	17	13	14	13	12	15	14	11	15	16	-	40	40	41	39	41	39	43
(12) *Chlamydia pneumoniae*	83	12	9	15	11	14	16	15	14	15	16	40	-	42	40	41	41	41	42
(13) *Bacillus stearothermophilus*	76	13	13	16	11	14	16	12	16	15	14	40	42	-	42	41	41	40	43
(14) *Escherichia coli* MalP	51	14	12	14	13	12	13	14	13	12	17	41	40	42	-	41	40	40	43
(15) *Escherichia coli* GlgP	60	11	12	16	11	16	17	13	14	16	15	39	41	41	41	-	41	40	43
(16) Potato type-H isozyme	53	12	13	15	10	13	14	13	12	14	15	41	41	41	40	41	-	41	43
(17) Potato type-L isozyme	57	12	14	13	9	12	12	11	11	13	17	39	41	40	40	40	41	-	43
(18) Rabbit muscle	66	15	10	16	12	15	15	16	15	17	15	43	42	43	43	43	43	43	-

Fig. 5. Comparison of amino acid sequences of glucan phosphorylases from various sources.

Amino acid sequences of 18 glucan phosphorylases were aligned using the CLUSTAL W (1.81) program (http://www2.ebi.ac.uk/clustalw/), and the alignment scores (%) were shown. Database accession numbers for each enzyme are, *Thermus aquaticus* (AB047267), *Deinococcus radioduruns* (AE 002052.1), *Thermococcus litoralis* (AF115479.1), *Streptomyces coelicolor* (AJ001205.2), *Pyrococcus abyssi* (AJ248285.1), *Pyrococcus horikoshii* (AP000006.1), *Mycobacterium tuberculosis* (Z73902.1), *Thermotoga maritima* (AJ001088.1), *Aquifex aeolicus* (AE000704.1), *Methanococcus jannaschii* (U 67603.1), *Pseudomonas aeruginosa* (AE004641.1), *Chlamydia pneumoniae* (AE001615.1), *Bacillus stearothermophilus* (D87026.1), *E. coli* malP (X06791.1), *E. coli* glgP (X16931.1), potato type-H (M 69038.1), potato type-L (D00520.1) and rabbit muscle (D00040.1).

similarity to the enzyme within this group, but lower similarity to others. Unfortunately, most of the enzymes within this new group are putative glucan phosphorylases and their properties have not been studied. However, two characterized enzymes (*T. aquaticus*, and *T. litoralis*) within this group, both have distinct substrate specificity as mentioned above. These results suggest the presence of a new sub-group in glucan phosphorylases which can be distinguished from the primary structure of the enzyme and by their substrate specificity.

This work was supported by a grant for the development of the next generation of bioreactor systems from the Society for Techno-Innovation of Agriculture, Forestry and Fisheries (STAFF).

REFERENCES

1) R. Shinzel and B. Nidetzky: Bacterial α-glucan phosphorylases. *FEMS Microbiol. Lett.*, **171**, 73–79 (1999).
2) K.B. Xavier, R. Peist, M. Kossmann, W. Boos and H. Santos: Maltose metabolism in the hyperthermophilic archaeon *Thermococcus litoralis*: Purification and characterization of key enzymes. *J. Bacteriol.*, **181**, 3358–3367 (1999).
3) D. Palm, R. Goerl and K.J. Burger: Evolution of catalytic and regulatory sites in phosphorylases. *Nature*, **313**, 500–502 (1985).
4) B. Boeck and R. Schinzel: Purification and characterization of an α-glucan phosphorylase from the thermophilic bacterium *Thermus thermophilus*. *Eur. J. Biochem.*, **239**, 150–155 (1996).
5) S. Kitamura: Starch polymers, natural and synthetic. in *Polymeric Materials Encyclopedia*, J.C. Salamone, Ed., CRC Press, Boca Raton, p. 7915 (1996).
6) A. Weinhausel, B. Nidetzky, C. Kysela and K.D. Kulbe: Application of *Escherichia coli* maltodextrin phosphorylase for the continuous production of glucose 1-phosphate. *Enzyme Microb. Technol.*, **17**, 140–146 (1995).
7) T. Takaha, M. Yanase, S. Okada and S.M. Smith: Disproportionating enzyme (4-α-glucanotransferase; EC 2.4.1.25) of potato. *J. Biol. Chem.*, **268**, 1391–1396 (1993).
8) S. Saheki, A. Takeda and T. Shimazu: Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. *Anal. Biochem.*, **148**, 277–281 (1985).
9) H. Takata, T. Takaha, S. Okada, M. Takagi and T. Imanaka: Purification and characterization of α-glucan phosphorylase from *Bacillus stearothermophilus*. *J. Ferment. Bioeng.*, **85**, 156–161 (1998).
10) Y. Terada, K. Fujii, T. Takaha and S. Okada: *Thermus aquaticus* ATCC 33923 amylomaltase gene clon-

ing and expression and enzyme characterization: Production of cycloamylose. *Appl. Environ. Microbiol.*, 65, 910–915 (1999).
11) Y. Nakamura and M. Imamura: Characterization of *Chlorella* phosphorylase: The glucan specificity and effect of temperature. *Phytochemistry*, 22, 2395–2399 (1983).

(Received May 22, 2000; Accepted September 29, 2000)

大腸菌で発現させた *Thermus aquaticus* 由来耐熱性ホスホリラーゼの構造と諸性質

鷹羽武史，柳瀬美千代，高田洋樹，岡田茂孝

江崎グリコ㈱生物化学研究所
(555-8502 大阪市西淀川区歌島 4-6-5)

Thermus aquaticus 由来耐熱性ホスホリラーゼの部分アミノ酸配列を利用して，本酵素遺伝子を単離し，塩基配列を決定した．本遺伝子のORF部分（2460塩基対）を *tac* プロモーター制御下に組み込んだ発現ベクターを保持する大腸菌は，耐熱性ホスホリラーゼ活性を示し，このことから本遺伝子が，ホスホリラーゼをコードしていることを証明した．大腸菌で発現させた本酵素をSDS-PAGEで単一バンドを示すまで精製し，諸性質を調べた．本酵素のグルカン合成反応における最小のプライマーはマルトトリオース，分解反応における最小の基質はマルトテトラオースであった．これらは通常のホスホリラーゼに比べ，それぞれグルコース1単位ずつ短く，新たな基質特異性を有する酵素であった．本酵素はその構造においても，通常のホスホリラーゼと異なり，新たなホスホリラーゼグループに属する酵素であると考えられた．

Exhibit 2

J. Mol. Biol. (1993) 234, 700–721

Evolution of Allosteric Control in Glycogen Phosphorylase

John W. Hudson[1], G. Brian Golding[2] and Michael M. Crerar[1]†

[1]*Department of Biology*
York University, North York,
Ontario, Canada, M3J 1P3

[2]*Department of Biology*
McMaster University, Hamilton,
Ontario, Canada, L8S 4K1

(Received 14 May 1993; accepted 30 June 1993)

In relation to the primary sequence and three-dimensional structure of rabbit muscle glycogen phosphorylase, we have carried out a comparative sequence analysis of phosphorylases from human, rat, *Dictyostelium*, yeast, potato and *Escherichia coli*. Based on sequence similarity, a large region of the protein is shared by these enzymes, extending from alpha-helix-1 to the last alpha-helix-33. Conserved residues are equally distributed between the N and C-terminal domains and occur primarily in buried residues. Phylogenetic analysis indicates that the two isozymes within either *E. coli*, potato or *Dictyostelium* are more closely related to each other than they are to other phosphorylases. Yeast phosphorylase is most closely related to the *Dictyostelium* isozymes. Mammalian muscle and brain isozymes are more closely related to each other than to the liver isozyme and the muscle isozyme is evolving at the slowest rate. All phosphorylases exhibit high conservation of active site and pyridoxal phosphate binding residues. Most phosphorylases also exhibit high conservation of sugar binding residues in the glycogen storage site. Phosphorylation and AMP binding site residues are poorly conserved in non-mammalian phosphorylases. In contrast, glucose-6-P binding residues are highly conserved in four of the seven non-mammalian enzymes. Analysis of interacting pairs of dimer contact residues indicates that they can be grouped into three relatively independent networks. One network contains phosphorylation and AMP binding residues and is poorly conserved in non-mammalian enzymes. A second network contains glucose-6-P binding residues and is highly conserved in enzymes containing a conserved glucose-6-P binding site. A third, conserved network contains residues within the tower helix and gate loop. A model for the evolution of allostery in phosphorylase is proposed, suggesting that glucose-6-P inhibition was an early control mechanism. The later creation of primarily distinct ligand binding sites for AMP/phosphorylation control may have allowed the establishment of a separate dimer contact network for propagating conformational changes leading to activation rather than inhibition of enzyme activity.

Keywords: glycogen phosphorylase; evolution; glucose-6-P control; dimer contact networks; phylogenetic relationships

1. Introduction

Alpha-glucan phosphorylase (EC 2.4.1.1) catalyzes the breakdown of storage polysaccharides, such as glycogen, into glucose-1-P and hence plays a central role in carbohydrate metabolism (for a recent review, see Newgard *et al.*, 1989). The enzyme has been examined from a variety of organisms including bacteria, yeast, slime mold, plants, insects, fish, amphibians and mammals. All phosphorylases require the cofactor, pyridoxal phosphate, for activity but differ in their affinity and specificity for polysaccharides, their modes of regulation and their physiological roles.

Rabbit muscle glycogen phosphorylase is the most extensively studied member of this family (Acharya *et al.*, 1991; Browner & Fletterick, 1992; Newgard *et al.*, 1989). The enzyme exists as a homodimer containing two identical subunits of

† Author to whom all correspondence should be addressed.

0022-2836/93/230700–22 $08.00/0

molecular mass 97,500. It is a complex allosteric enzyme that is subject to both positive and negative control. Intracellular ligands, such as AMP and glycogen, activate the enzyme by promoting formation of the active R conformer, while glucose, glucose-6-P and purine nucleosides are thought to inhibit activity by stabilizing the inactive T conformer. Detailed X-ray crystallographic analysis of the three-dimensional structure of the enzyme (Acharya *et al.*, 1991; Barford *et al.*, 1991; Sprang *et al.*, 1991; Johnson *et al.*, 1992) indicates that AMP, glycogen, glucose and purine nucleosides bind to distinct sites on the enzyme: the AMP activation site, glycogen storage site, active site and purine nucleoside inhibitor site, respectively. Glucose-6-P binds in a spatial volume that overlaps with that used by AMP (Sprang *et al.*, 1988). However, most of the amino acid residues that interact with glucose-6-P are distinct from those that interact with AMP (Johnson *et al.*, 1992; Lorek *et al.*, 1984; Sprang *et al.*, 1987, 1988, 1991).

Covalent phosphorylation at Ser14 also promotes enzyme activation and this, in turn, is under extracellular control by factors that regulate phosphorylase kinase and phosphatase activity (Newgard *et al.*, 1989). In the three-dimensional structure of the enzyme, Ser14-P is situated close to AMP and it is thought that phosphorylation and AMP activate the enzyme by a similar overall mechanism but with local structural differences (Barford & Johnson, 1989; Barford *et al.*, 1991; Johnson *et al.*, 1992; Sprang *et al.*, 1987; 1991).

Based on the three-dimensional structure and positions of regulatory and active site residues, the enzyme monomer can be divided into two domains, the N-terminal and C-terminal domains (Browner & Fletterick, 1992; Newgard *et al.*, 1989). The N-terminal domain extends from amino acid residue (aa†) 1 to aa482 and is referred to as the "regulatory" domain since it contains the majority of ligand binding residues. It also contains all but one of the residues that form contacts with residues in the other, diad-related subunit (dimer contact residues). These residues are required for transmitting conformational changes across subunits upon ligand binding. Within the N-terminal domain, an "activation subdomain" exists, from aa1 to 120 (Goldsmith *et al.*, 1989*b*; Sprang *et al.*, 1991), which has been shown to undergo complex tertiary and quaternary changes on ligand binding (Browner *et al.*, 1992*a*). The C-terminal domain (aa483 to 842) is referred to as the "catalytic" domain since it contains the majority of active site and pryidoxal phosphate binding residues. The active site cleft is located between the N and C-terminal domains.

Three phosphorylase isozymes exist in mammals, muscle (M), brain (B) and liver (L), which derive their names from the tissues where they predominate (Newgard *et al.*, 1989). All three isozymes are structurally related but are encoded by distinct genes located on separate chromosomes in mouse (Glaser *et al.*, 1989) and human (Newgard *et al.*, 1989). Interestingly, the three isozymes exhibit differences in their ability to be fully activated by either phosphorylation or AMP in accord with their different physiological roles. Comparative analysis of the amino acid sequences of these isozymes has led to the identification of potentially important, isozyme specific substitutions that may play a role in conferring their differential responses to allosteric control (Hudson *et al.*, 1993; Rath *et al.*, 1987). Analysis of engineered single and multiple site exchanges of these substitutions between isozymes is now being performed (Browner *et al.*, 1992*b*; Coats *et al.*, 1991).

In lower organisms, phosphorylases vary dramatically in their ability to respond to allosteric control mechanisms. For example, bacterial and plant phosphorylases are active in the absence of phosphorylation and AMP (Newgard *et al.*, 1989). The yeast enzyme requires phosphorylation for maximal activity, but is insensitive to activation by AMP (Rath, 1991; Rath *et al.*, 1992). Phosphorylation occurs at a different residue (Thr) and likely by a different mechanism than that used for activation of the mammalian isozymes. Phosphorylase 1 from *Dictyostelium discoideum* can be activated by either phosphorylation or AMP (Naranan *et al.*, 1988; Rutherford & Cloutier, 1986). The phosphorylated residue in this enzyme has not been identified and it is unknown if the mechanism of phosphorylation is similar to that used by yeast or mammals. A second genetically distinct phosphorylase in *Dictyostelium*, phosphorylase 2, does not require phosphorylation or AMP for activity (Rutherford *et al.*, 1992). However, this enzyme may initially exist in an inactive form and it has been postulated that proteolytic cleavage of N-terminal residues may be necessary for activation. With respect to the negative control of these enzymes, glucose-6-P has been shown to inhibit the activity of the yeast enzyme (Sagardia *et al.*, 1971).

Non-mammalian phosphorylases also exhibit wide differences in their polysaccharide affinities and specificites (Newgard *et al.*, 1989). For example, *Escherichia coli* contains both a true glycogen phosphorylase and a maltodextrin phosphorylase that shows specificity for low molecular mass, unbranched alpha-1.4 polyglucoses. Glycogen and maltodextrin phosphorylase in *E. coli* are encoded by distinct genes, *glg* P and *mal* P present in the glycogen and maltose regulons, respectively, and maltodextrin phosphorylase is inducible by maltose (Choi *et al.*, 1989; Newgard *et al.*, 1989; Palm *et al.*, 1987; Yu *et al.*, 1988). Two genetically distinct phosphorylases also exist in potato tubers and spinach leaves (Mori *et al.*, 1991; Nakano *et al.*, 1989). One of these, type H, is present in the cytosol and has high affinity for a variety of glucans including glycogen. The other phosphorylase, type L, which is localized in plastids, has very low affinity for glycogen but high affinity for amylose, amylopectin and malto-

† Abbreviations used: aa, amino acid residue; M, B and L isozymes, isozymes derived from muscle, brain and liver respectively.

dextrin (Fukui *et al.*, 1982; Shimomura & Fukui, 1980; Shimomura *et al.*, 1982).

As these enzymes exhibit dramatic differences in their responses to allosteric control mechanisms and in their polysaccharide specificities, phosphorylase is an attractive enzyme for the study of the evolution of enzyme allostery and substrate specificity. Since the three-dimensional structure of the rabbit M isozyme has been solved to high resolution, amino acid sequence comparisons of these enzymes can be made in relation to the primary sequence and three-dimensional structure and function of the rabbit M isozyme. Complete amino acid sequences have been determined for 14 phosphorylases. These include the M isozyme from rabbit (Nakano *et al.*, 1986), rat (Hudson *et al.*, 1993) and human (Burke *et al.*, 1987), the B isozyme from rat (Hudson *et al.*, 1993) and human (Gelinas *et al.*, 1989; Newgard *et al.*, 1988), the L isozyme from rat (Scheibel *et al.*, 1992) and human (Newgard *et al.*, 1986), phosphorylase 1 (Rogers *et al.*, 1992) and phosphorylase 2 (Rutherford *et al.*, 1992) from *Dictyostelium*, yeast phosphorylase (Hwang & Fletterick, 1986; Rath *et al.*, 1992), type H (Mori *et al.*, 1991) and type L (Nakano *et al.*, 1989) phosphorylases from potato and *E. coli* glycogen (Choi *et al.*, 1989; Yu *et al.*, 1988) and maltodextrin (Palm *et al.*, 1985, 1987) phosphorylase.

In the past, comparative sequence analyses carried out by ourselves and others have been limited to the study of subsets of these sequences (Crerar *et al.*, 1988; Choi *et al.*, 1989; Hwang & Fletterick, 1986; Mori *et al.*, 1991; Nakano & Fukui, 1986; Newgard *et al.*, 1989; Palm *et al.*, 1985; Rogers *et al.*, 1992; Rutherford *et al.*, 1992; Yu *et al.*, 1988). To obtain a more complete evolutionary analysis of phosphorylase, we present in this report, a comparison of the amino acid sequences of all of these enzymes in relation to recent findings concerning the three-dimensional structure and function of the rabbit M isozyme (Barford *et al.*, 1991; Goldsmith *et al.*, 1989*a*; Johnson *et al.*, 1990, 1992; Sprang *et al.*, 1991).

2. Materials and Methods

(a) *Phosphorylase sequences*

The complete amino acid sequences of 14 phosphorylases were used for comparative sequence analysis. These were obtained from the following published sources: rabbit M isozyme (Nakano *et al.*, 1986), human M isozyme (Burke *et al.*, 1987), rat M isozyme (Hudson *et al.*, 1993), human B isozyme (Newgard *et al.*, 1988, Gelinas *et al.*, 1989), rat B isozyme (Hudson *et al.*, 1993), human L isozyme (Newgard *et al.*, 1986), rat L isozyme (Scheibel *et al.*, 1992), phosphorylase 1 from *Dictyoselium* (Rogers *et al.*, 1992), phosphoprylase 2 from *Dictyostelium* (Rutherford *et al.*, 1992), yeast phosphorylase (Hwang & Fletterick, 1986; Rath *et al.*, 1992), type H phosphorylase from potato (Mori *et al.*, 1991), type L phosphorylase from potato (Nakano *et al.*, 1989), *E. coli* glycogen phosphorylase (Choi *et al.*, 1989; Yu *et al.*, 1988) and *E. coli* maltodextrin phosphorylase (Palm *et al.*, 1985, 1987). For yeast phosphorylase, minor corrections of the published sequence were communicated to us by P. K. Hwang and R. J. Fletterick. These occur at aa267, aa456, aa472 and aa805.

(b) *Alignment of phosphorylase sequences*

Phosphorylase sequences were initially aligned using the programme CLUSTAL V (Higgins & Sharp, 1989). A number of different combinations of gap penalties and different sets of sequences were used to obtain an alignment. Additionally, some regions were aligned by eye to maximize sequence similarity and to minimize the number of small deletions/insertions.

(c) *Phylogenetic relationships of phosphorylases*

The phylogenetic relationships of these phosphorylases were determined using the neighbor joining method of Saitou & Nei (1987). Trees were bootstrapped 1000 times by sampling sites at random with replacement. These samples were used to generate new sequences or distance matrices in order to reconstruct new phylogenies. A consensus tree was obtained from the bootstrapped samples. For the purpose of illustration, the trees were arbitrarily rooted at the longest branch and trees were drawn using the computer program, Canvas version 3.0. In addition to the neighbor joining method, phylogenies were also reconstructed using the protein parsimony method (Felsenstein, 1990). Phylogenetic trees with similar topologies were obtained with both methods. Strict consensus sequences for the overall and mammalian ancestral enzymes were determined using the program, Phylipp version 3.3 (Felsenstein, 1990).

(d) *Rates of change*

Rates of change of residues throughout phosphorylase were determined using a program designed by G. B. Golding (personnal communication).

3. Results

(a) *Conservation of the overall structure of phosphorylase*

An alignment of the complete amino acid sequences of all 14 phosphorylases is presented in Figure 1. The numbering of amino acid residues corresponds to that used for the rabbit M isozyme. Insertions and deletions were placed in the sequences to maximize similarity between all 14 enzymes. As noted previously (Mori *et al.*, 1991; Nakano *et al.*, 1989; Newgard *et al.*, 1989; Rogers *et al.*, 1992; Rutherford *et al.*, 1992), there is considerable variation in the lengths of these enzymes at their N and C-termini. However, in relation to the primary sequence and three-dimensional structure of the rabbit M isozyme, sequence conservation appears to begin close to the N terminus, at helix-1 (aa23 to 40) and extends to at least the end of the last alpha-helix, helix-33, (aa813 to 825).

The degree of conservation of the overall structure of the enzyme was determined in relation to aa17 to 829 of the rabbit M isozyme (see Table 1). This corresponds to the relative positions in the alignment of the N and C-termini of the shortest phosphorylase, *E. coli* maltodextrin phosphorylase.

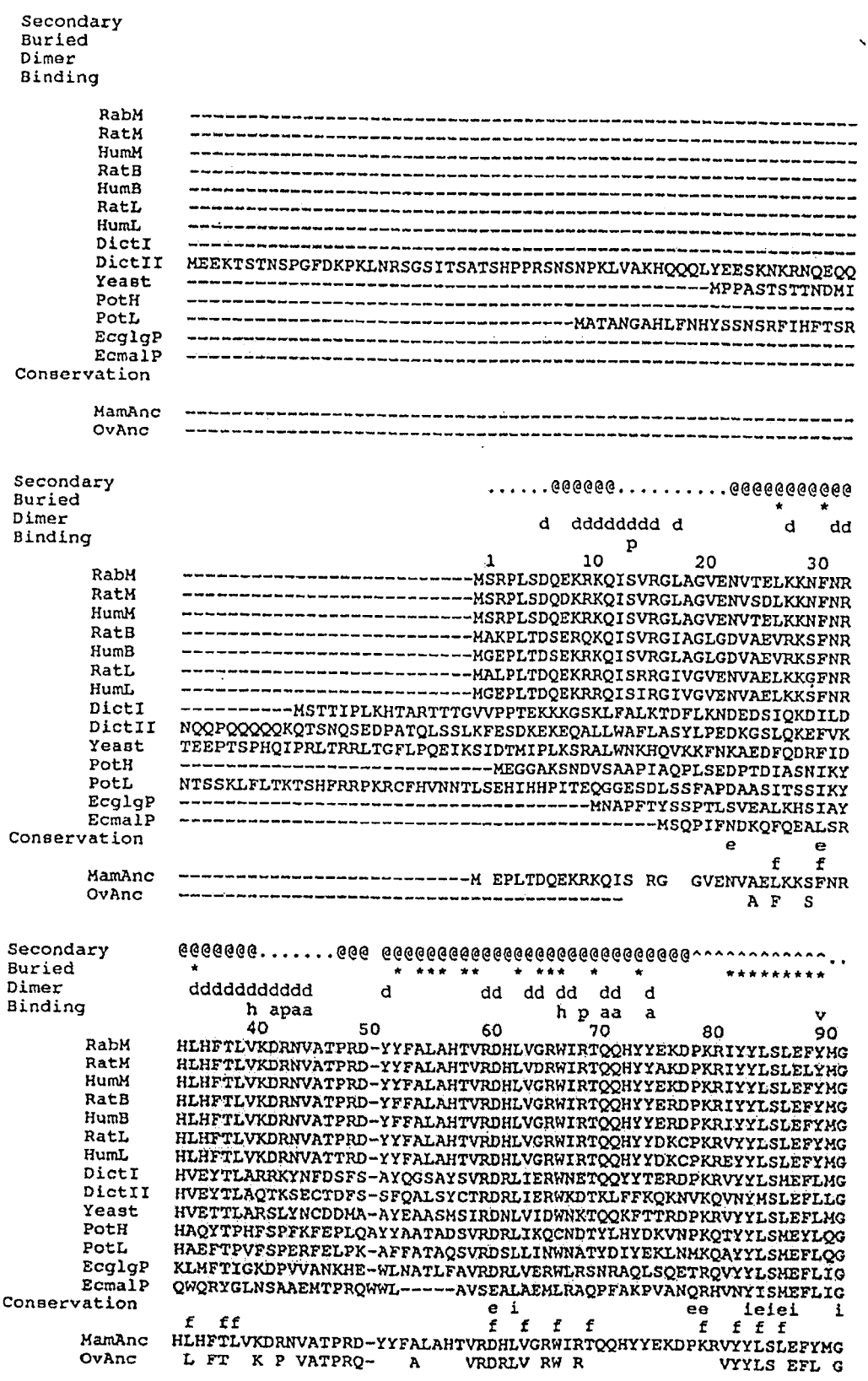

Secondary
Buried
Dimer
Binding
RabM
RatM
HumM
RatB
HumB
RatL
HumL
DictI
DictII
Yeast
PotH
PotL
EcglgP
EcmalP
Conservation
MamAnc
OvAnc
1 10 20 30
40 50 60 70 80 90

Fig. 1.

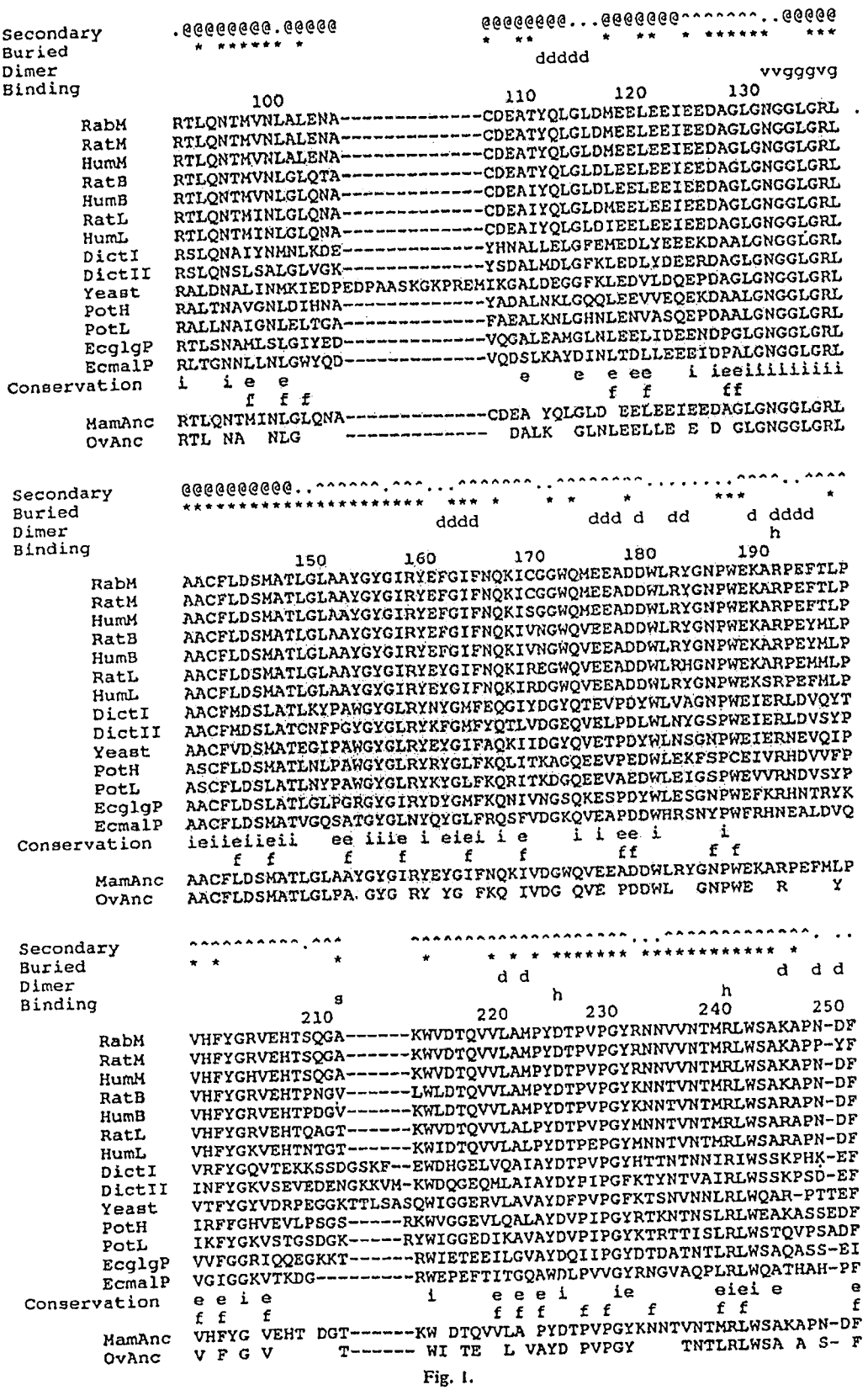

Secondary
Buried
Dimer
Binding
100
110
120
130
RabM RTLQNTMVNLALENA----------CDEATYQLGLDMEELEEIEEDAGLGNGGLGRL
RatM RTLQNTMVNLALENA----------CDEATYQLGLDMEELEEIEEDAGLGNGGLGRL
HumM RTLQNTMVNLALENA----------CDEATYQLGLDMEELEEIEEDAGLGNGGLGRL
RatB RTLQNTMVNLGLQTA----------CDEATYQLGLDLEELEEIEEDAGLGNGGLGRL
HumB RTLQNTMVNLGLQNA----------CDEAIYQLGLDLEELEEIEEDAGLGNGGLGRL
RatL RTLQNTMINLGLQNA----------CDEAIYQLGLDMEELEEIEEDAGLGNGGLGRL
HumL RTLQNTMINLGLQNA----------CDEAIYQLGLDIEELEEIEEDAGLGNGGLGRL
DictI RSLQNAIYNMNLKDE----------YHNALLELGFEMEDLYEEEKDAALGNGGLGRL
DictII RSLQNSLSALGLVGK----------YSDALMDLGFKLEDLYDEERDAGLGNGGLGRL
Yeast RALDNALINMKIEDPEDPAASKGKPREMIKGALDEGGFKLEDVLDQEPDAGLGNGGLGRL
PotH RALTNAVGNLDIHNA----------YADALNKLGQQLEEVVEQEKDAALGNGGLGRL
PotL RALLNAIGNLELTGA----------FAEALKNLGHNLENVASQEPDAALGNGGLGRL
EcglgP RTLSNAMLSLGIYED----------VQGALEAMGLNLEELIDEENDPGLGNGGLGRL
EcmalP RLTGNNLLNLGWYQD----------VQDSLKAYDINLTDLLEEEIDPALGNGGLGRL
Conservation
MamAnc RTLQNTMINLGLQNA----------CDEA YQLGLD EELEEIEEDAGLGNGGLGRL
OvAnc RTL NA NLG ---------- DALK GLNLEELLE E D GLGNGGLGRL
Secondary
Buried
Dimer
Binding
150
160
170
180
190
RabM AACFLDSMATLGLAAYGYGIRYEFGIFNQKICGGWQMEEADDWLRYGNPWEKARPEFTLP
RatM AACFLDSMATLGLAAYGYGIRYEFGIFNQKICGGWQMEEADDWLRYGNPWEKARPEFTLP
HumM AACFLDSMATLGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEYMLP
RatB AACFLDSMATLGLAAYGYGIRYEFGIFNQKIVNGWQVEEADDWLRYGNPWEKARPEYMLP
HumB AACFLDSMATLGLAAYGYGIRYEFGIFNQKIVNGWQVEEADDWLRYGNPWEKARPEYMLP
RatL AACFLDSMATLGLAAYGYGIRYEYGIFNQKIREGWQVEEADDWLRHGNPWEKARPEMMLP
HumL AACFLDSMATLGLAAYGYGIRYEYGIFNQKIRDGWQVEEADDWLRYGNPWEKSRPEFMLP
DictI AACFMDSLATLKYPAWGYGLRYNYGMFEQGIYDGYQTEVPDYWLVAGNPWEIERLDVQYT
DictII AACFMDSLATCNFPGYGYGLRYKFGMFYQTLVDGEQVELPDLWLNYGSPWEIERLDVSYP
Yeast AACFVDSMATEGIPAWGYGLRYEYGIFAQKIIDGYQVETPDYWLNSGNPWEIERNEVQIP
PotH ASCFLDSMATLNLPAWGYGLRYRYGLFKQLITKAGQEEVPEDWLEKFSPCEIVRHDVVFP
PotL ASCFLDSLATLNYPAWGYGLRYKYGLFKQRITKDGQEEVAEDWLEIGSPWEVVRNDVSYP
EcglgP AACFLDSLATLGLPGRGYGIRYDYGMFKQNIVNGSQKESPDYWLESGNPWEFKRHNTRYK
EcmalP AACFLDSMATVGQSATGYGLNYQYGLFRQSFVDGKQVEAPDDWHRSNYPWFRHNEALDVQ
Conservation
MamAnc AACFLDSMATLGLAAYGYGIRYEYGIFNQKIVDGWQVEEADDWLRYGNPWEKARPEFMLP
OvAnc AACFLDSMATLGLPA GYG RY YG FKQ IVDG QVE PDDWL GNPWE R Y
Secondary
Buried
Dimer
Binding
210
220
230
240
250
RabM VHFYGRVEHTSQGA------KWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAKAPN-DF
RatM VHFYGRVEHTSQGA------KWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAKAPP-YF
HumM VHFYGHVEHTSQGA------KWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAKAPN-DF
RatB VHFYGRVEHTPNGV------LWLDTQVVLAMPYDTPVPGYKNNTVNTMRLWSAKAPN-DF
HumB VHFYGRVEHTPDGV------KWLDTQVVLAMPYDTPVPGYKNNTVNTMRLWSARAPN-DF
RatL VHFYGRVEHTQAGT------KWVDTQVVLALPYDTPVPGYMNNTVNTMRLWSARAPN-DF
HumL VHFYGKVEHTNTGT------KWIDTQVVLALPYDTPEPGYMNNTVNTMRLWSARAPN-DF
DictI VRFYGQVTEKKSSDGSKF--EWDHGELVQAIAYDTPVPGYHTTNTNNIRIWSSKPHK-EF
DictII INFYGKVSEVEDENGKKVM-KWDQGEQMLAIAYDYPIPGFKTYNTVAIRLWSSKPSD-EF
Yeast VTFYGYVDRPEGGKTTLSASQWIGGERVLAVAYDFPVPGFKTSNVNNLRLWQAR-PTTEF
PotH IRFFGHVEVLPSGS------RKWVGGEVLQALAYDVPIPGYRTKNTNSLRLWEAKASSEDF
PotL IKFYGKVSTGSDGK------RYWIGGEDIKAVAYDVPIPGYKTRTTISLRLWSTQVPSADF
EcglgP VVFGGRIQQEGKKT------RWIETEEILGVAYDQIIPGYDTDATNTLRLWSAQASS-EI
EcmalP VGIGGKVTKDG---------RWEPEFTITGQAWDLPVVGYRNGVAQPLRLWQATHAH-PF
Conservation
MamAnc VHFYG VEHT DGT------KW DTQVVLA PYDTPVPGYKNNTVNTMRLWSAKAPN-DF
OvAnc V F G V T------ WI TE L VAYD PVPGY TNTLRLWSA A S- F

Fig. 1.

```
Secondary       .........@@@@@@@@@@@@@@@@^^^^^^.^^^^^@@@@@@@@@@@@@@@@@@@@@@@
Buried                                **  *  **        ** ********** ***
Dimer            d       dd  dd dd   d  ddddd         d
Binding                                      cggc                       aa
                       260       270       280       290       300       310
RabM            NLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
RatM            NLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
HumM            NLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
RatB            KLKDFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
HumB            KLQDFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVGATLQDIIRRFK
RatL            NLQDFNVGDYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDVIRRFK
HumL            NLRDFNVGDYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
DictI           DLDAFNGGNYLSAVEAKQRSENITSVLYPNDNTYSGKELRLKQQYFFVAATLCDVIRRFK
DictII          NLDSFNRGDYLGAIEEKEKSENITNVLYPNDNTMQGKELRLKQQYLFVSATIQDIISQFK
Yeast           DFAKFNNGDYKNSVAQQQRAESITAVLYPNDNFAQGKELRLKQQYFWCAASLHDILRRFK
PotH            NLFLFNDGQYDAAAQLHSRAQQICAVLYPGDATENGKLLRLKQQFFLCSASLQDIIARFK
PotL            DLSAFNAGEHTKACEAQANAEKICYILYPGDESEEGKILRLKQQYTLCSASLQDIISRFE
EcglgP          NLGKFNQGDYFAAVEDKNHSENVSRVLYPDDSTYSGRELRLRQEYFLVSSTIQDILSR--
EcmalP          DLTKFNDGDFLRAEQQGINAEKLTKVLYPNDNHTAGKKLRLMQQYFQCACSVADILRR--
Conservation     e  ii i    e      ee e  eiii i    ie iii iee  e ee iee e
                 f                    f  f         f           f ff  ff
MamAnc          NL DFNVGDYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFK
OvAnc            L  FNDGDYLAAVED N AEN  RVLYPNDNTY GKELRL Q YFL A   QDI RR

Secondary       .............@@@@@@@^^^^^^....@@@@@@@@@@@@@@@...@@@@@@@@@@@^
Buried                       *  **  **********************         *  **  **
Dimer
Binding           aa a                                   ss  s ss
                       320       330       340       350       360       370
RabM            SSKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPELMRVLVDLERLDWDKAWEVTVKTC
RatM            SSKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPELIRILVDLERLDWDKAWDVTVKTC
HumM            SSKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPELMRILVDLERMDWDKAWDVTVRTC
RatB            SSKFGCRDPVRTCFETFPDKVAIQLNDTHPALSIPELMRILVDVEKVDWDKAWEITKKTC
HumB            SSKFGCRDPVRTCFETFPDKVAIQLNDTHPALSIPELMRILVDVEKVDWDKAWEITKKTC
RatL            ASKFGSKDGVGTVFDAFPDQVAIQLNDTHPALAIPELMRIFVDIEKLPWSKAWEITKKTF
HumL            ASKFGSTRGQGTVFDAFPDQVAIQLNDTHPRIAIPELMRIFVDIEKLPWSKAWELNQKTF
DictI           KS--------HQNWQDFPNKVAIQLNDTHPTIGVVELFRKLIDEEGLQWEEAWDIVTKTF
DictII          ET--------GKPFSEFTFH-AIQLNDTHPTLGIPELMRILIDEEKKSWDEAWDITTKTF
Yeast           KSK--------RPWTEFPDQVAIQLNDTHPTLAIVELQRVLVDLEKLDWHEAWDIVTKTF
PotH            ERE---DGKGSHQWSEFPKKVAIQLNDTHPTLTIPELMRLLMDDEGLGWDESWNITTRTI
PotL            RRS-----GDRIKWEEFPEKVAVQMNDTHPTLCIPELMRILIDLKGLNWNEAWNITQRTV
EcglgP          ------HYQLHKTYDNLADKIAIHLNDTHPVLSIPEMMRLLIDEHQFSWDDAFEVCCQVF
EcmalP          ------HHLAGRELHELADYEVIQLNDTHPTIAIPELLRVLIDEHQMSWDDAWAITSKTF
Conservation                    e     eeeiiiii e e ie i eei     i  e  e  e
                                f    ff f      f ff f   ff          f f    f
MamAnc           SKFG  DGVGT FDAFPDKVAIQLNDTHPALAIPELMRILVDLEKLDWDKAWEITKKTF
OvAnc              ---    G   DE  DK AIQLNDTHPTLAIPELMR LID     WD AWEIT KTF

Secondary       ^^^^^^.....^^^^^@@@@@@@.@@@@@@@@@@@@@@@@@@@@@...
Buried          **** * *  * * *** **  *** ** **  **  **  *
Dimer
Binding             g                    s  ss ss sss s
                         380       390       400       410       420
RabM            AYTNHTVLPEALERWPVHLLETLLPRHLQIIYEINQRFLNRVAAAFPG------------
RatM            AYTNHTVLPEALERWPVHLMETLLPRHLQIIYEINQRFLNRVAAAFPG------------
HumM            AYTNHTVLPEALERWPVHLLETLLPRHLQIIYEINQRFLNRVAAAFPG------------
RatB            AYTNHTVLPEALERWPVSMFEKLLPRHLEIIYAINQRHLDHVAALFPG------------
HumB            AYTNHTVLPEALERWPVSMFEKLLPRHLEIIYAINQRHLDHVAALFPG------------
RatL            AYTNHTVLPEALERWPVDLVEKLLPRHLQIIYEINQKHLDRIVALFPK------------
HumL            AYTNHTVLPEALERWPVDLVEKLLPRHLEIIYEINQKHLDRIVALFPK------------
DictI           AYTNHTILPEALEMWPVSLIEDLLPRHMQLIYGINHRFLIQVTQKWPG------------
DictII          SYTNHTVLPEALEKWSVSHVENVLPRHIMIIYEINERFLKLVDQKWPG------------
Yeast           AYTNHTVMQEALEKWPRRLFGHLLPRHLEIIYDINWFFLQDVAKKFPK------------
PotH            AYTNHTVLPEALEKWSQAVMWKLLPRHMEIIEEIDKRFVATIMSERPD------------
PotL            AYTNHTVLPEALEKWSYELMQKLLPRHVEIIEAIDEELVHEIVLKYGSMDLNKLEEKLTT
EcglgP          SYTNHTLMSEALETWPVDMLGKILPRHLQIIFEINDYFLKTLQEQYPN------------
EcmalP          AYTNHTLMPEALERWDVKLVKGLLPRHMQIINEINTRFKTLVEKTWPG------------
Conservation    eiiiiiee iiii i   e   eiiiie ei  ie      e
                      ff          ff  f    f f           f    f
MamAnc          AYTNHTVLPEALERWPVSLVEKLLPRHL IIYEINQR LDRV ALFPG------------
OvAnc           AYTNHT  PEALERWPV LV KLLPRH  II EIN RFLK V K  PG------------
```

Fig. 1.

```
Secondary
Buried
Dimer
Binding
         RabM   ------------------------------------------------------------
         RatM   ------------------------------------------------------------
         HumM   ------------------------------------------------------------
         RatB   ------------------------------------------------------------
         HumB   ------------------------------------------------------------
         RatL   ------------------------------------------------------------
         HumL   ------------------------------------------------------------
        DictI   ------------------------------------------------------------
       DictII   ------------------------------------------------------------
        Yeast   ------------------------------------------------------------
         PotH   ------------------------------------------------------------
         PotL   MRILENFDLPSSVAELFIKPEISVDDDTETVEVHDKVEASDKVVTNDEDDTGKKTSVKIE
       EcglgP   ------------------------------------------------------------
       EcmalP   ------------------------------------------------------------
Conservation

       MamAnc   ------------------------------------------------------------
        OvAnc   ------------------------------------------------------------


Secondary            .@@@@@@@^^^^^ ..^^^^^@@@@@@@...^^^^^^@@@@@@@@..@@@@@@@@
Buried                 *   ****        *********** ******  **   *     * **
Dimer
Binding               ss ss  s s s   s               g
                              430       440       450       460       470
         RabM   -----DVDRLRRMSLVEE-GAVKRINMAHLCIAGSHAVNGVARIHSEILKKTIFKDFYEL
         RatM   -----DVDRLRRMSLVEE-GAVKRINMAHLCIAGSHAVNGVARIHSEILKKTIFKDFYEL
         HumM   -----DVDRLRRMSLVEE-GAVKRINMAHLCIAGSHAVNGVARIHSEILKKTIFKDFYEL
         RatB   -----DVDRLRRMSVIEE-GDCKRINMAHLCVIGSHAVNGVARIHSEIVKQSVFKDFYEL
         HumB   -----DVDRLRRMSVIEE-GDCKRINMAHLCVIGSHAVNGVARIHSEIVKQSVFKDFYEL
         RatL   -----DIDRMRRMSLIEE-EGGKRINMAHLCIVGCHAVNGVAKIHSDIVKTQVFKDFSEL
         HumL   -----DVDPLRRMSLIEE-EGSKRINMAHLCIVGSHAVNGVAKIHSDIVKTKVFKDFSEL
        DictI   -----DIGKMRGLSIIQE-GEEKRVRMAHLAIVGSHCVNGVAAMHSELVKHKVFPDFFCL
       DictII   -----DMSKRRALSIIDE-SDGKFIVMAFLAIVGAHTINGVAYLHSELVKHDVFPLFYEV
        Yeast   -----DVDLLSRISIIEENSPERQIRMAFLAIVGSHKVNGVAELHSELIKTTIFKDFVKF
         PotH   -----LENKMPSMRILDHNATKPVVHMANLCVVSSHTVNGVAQLHSDILKAELFADYVSV
         PotL   AAAEKDIDKKTPVSPEPAVIPPKKVRMANLCVVAGHAVNGVAEIHSEIVKEEVFNDFYEL
       EcglgP   -----DTDLLGRASIIDE-SNGRRVRMAWLAVVVSHKVNGVSELHSNLMVQSLFADFAKI
       EcmalP   -----DEKVWAKLAVVHD----KQVHMANLCVVGGFAVNGVAALHSDLVVKDLFPEYHQL
Conservation                          e ii i e     eiiie eiieee   ei  e
                             f f      f    fff     f    f  ff    f      f
       MamAnc   -----D DRLRRMSLIEE-  GKRINMAHLCIVGSHAVNGVA IHSEIVK  VFKDFYEL
        OvAnc   -----D D L R SIIDE-S GKR RMANLC VGSHAVNGVAELHS  V     F DF  L


Secondary       . @@@^^^^^^^^^^@@@@@@@@.@@@@@@@@@@..@    @@@..@@@@@@@@@....@@@
Buried          *   **  ***************   ** **          **  *     *   *    *
Dimer
Binding                   g      v
                    480       490       500           510      520
         RabM   E-PHKFQNKTNGITPRRWLVLCNPGLAEIIAERIGE---EYISDLDQLRKLLSYVDDEAF
         RatM   E-PHKFQNKTNGITPRRWLVLCNPGLAEVIAERIGE---EYISDLDQLRKLLSYLDDQAF
         HumM   E-PHKFQNKTNGITPRRWLVLCNPGLAEVIAERIGE---DFISDLDQLRKLLSFVDDEAF
         RatB   E-PEKFQNKTNGITPRRWLLLCNPGLAEIIVERIGE---GFLTDLSQLKKLLSLVDDEAF
         HumB   E-PEKFQNKTNGITPRRWLLLCNPGLADTIVEKIGE---EFLTDLSQLKKLLPLVSDEVF
         RatL   E-PDKFQNKTNGITPRRWLLLCNPGLADLIAEKIGE---DYVKDLSQLTKLHSFVGDDIF
         HumL   E-PDKFQNKTNGITPRRWLLLCNPGLAELIAEKIGE---DYVKDLSQLTKLHSFLGDDVF
        DictI   W-PEKFQNKTNGVTPRRWIEQANPGLSAIFTKWLGT--DKWTTNLELVKGIKEHMDNPEL
       DictII   W-PNKFQSKTSGVTPSSWIEQSNPQLAELITRSLNS--DRWLVNLDIIKDLVHLADNSSF
        Yeast   YGPSKFVNVTNGITPRRWLKQANPSLAKLISETLNDPTEEYLLDMAKLTQLEKYVEDKEF
         PotH   W-PTKFQNKTNGITPRRWIRFCSPELSHIITKWLKT--DQWVTNLELLANLREFADNSEL
         PotL   W-PEKFQNKTNGVTPRRWIRFCNPPLSAIITKWTGT--EDWVLKTEKLAELQKFADNEDL
       EcglgP   F-PGRFTNVTNGVTPRRWLAVANPSLSAVLDGTLGR---NWRTDLSLLNELQQHCDFPMV
       EcmalP   W-PNKFHNVTNGITPRRWIKQCNPALAALLDKSLQK---EWANDLDQLINLVKLADDAKF
                                                            e  e
Conservation      i ei e ieieii  ie   ei ie  e              f  f    f    f
                    f  f  f f     f    f     f
       MamAnc   E-PEKFQNKTNGITPRRWLLLCNPGLAELIAE IGE---EY TDLSQL KL SLVDD  F
        OvAnc   W-P KF N TNGITPRRW KQCNP LAAL  K LG ---EW TDLS L  L KLADD  F
```

Fig. 1.

```
Secondary      @@@@@@@@@@@@@@@@@@@@@@...      .....^^^^^^^^.......@@@@@@@@@
Buried         *  *  **  **  **  *             *   ************************
Dimer
Binding                                                  vvv c gg
                        540       550             560       570        580
      RabM     IRDVAKVKQENKLKFAAYLEREYKV------HINPNSLFDVQVKRIHEYKRQLLNCLHVI
      RatM     IRDVAKVKQENKLKFSAYLETEYKV------HINPNSLFDVQVKRIHEYKRQLLNCLHII
      HumM     IRDVAKVKQENKLKFAAYLEREYKV------HINPNSLFDIQVKRIHEYKRQLLNCLHVI
      RatB     IRDVAKVKQENKLKFSAQLEKEYKV------KINPASMFDVHVKRIHEYKRQLLNCLHII
      HumB     IRDVAKVKQENKLKFSAFLEKEYKV------KINPSSMFDVHVKRIHEYKRQLLNCLHVV
      RatL     LREIAKVKQENKLKFSQFLEKEYKV------KINPSSMFDVHVKRIHEYKRQLLNCLHVI
      HumL     LRELAKVKQENKLKFSQFLETEYKV------KINPSSMFDVQVKRIHEYKRQLLNCLHVI
      DictI    IAEWKYVKQGNKQRLAEFILKHCGI------HVNPNALFDVHIKRIHEYKRQLLNILSVI
      DictII   QKEWMTINRNNKIRLAKYIEKRCDI------QVNVDVLFDVQVKRFHEYKRQLLNVLSVI
      Yeast    LKKWNQVKLNNKIRLVDLIKKENDGVDIINREYLDDTLFDMQVKRIHEYKRQQLNVFGII
      PotH     HAEWESAKMANKQRLAQYILHVTGV------SIDPNSLFDIQVKRIHEYKRQLLNILGVI
      PotL     QNEWREAKRSNKIKVVSFLKEKTGY------SVVPDAMFDIQVKRIHEYKRQLLNIFGIV
      EcglgP   NHAVHQANVENKKRLAEYIAQQLNV------VVNPKALFDVQIKRIHEYKRQLMNVLHVI
      EcmalP   RDLYRVIKQANKVRLAEFVKVRTGI------DINPQAIFDIQIKRLHEYKRQHLNLLHIL
Conservation          e ii  e    e                  eiieeeiieiiiii ei e ee
                     f     ff    f                  f  f f  f       f ff ff
      MamAnc    REVAKVKQENKLKFSQFLEKEYKV------KINPSS FDVQVKRIHEYKRQLLNCLHVI
      OvAnc        VRQ KQ NK RLA FIK  TGV------  NP ALFDVQ KRIHEYKRQLLNVLHVI

Secondary      @@@@@@@@         ..^^^^^^^^^^^^^^^^...@@@@@@@@@@@@@@@@@@....
Buried         **** *                 * **********     * *** ** ***   *
Dimer                                             d
Binding                                             c
                    590                600       610       620       630
      RabM     TLYNRIKK---------EPNKFVVPRTVHIGGKAAPGYHMAKMIIKLITAIGDVVNHDPV
      RatM     TLYNRIKR---------EPNRFMVPRTIMIGGKAAPGYHMAKMIIKLITAIGDVVNHDPA
      HumM     TLYNRIKR---------EPNKFFVPRTVHIGGKAAPGYHMAKMIIRLVTAIGDVVNHDPA
      RatB     TLYNRIKK---------DPTKTFVPRTVMIGGKAAPGYHMAKMIIKLVTSIGDVVNHDPV
      HumB     TLYNRIKR---------DPAKAFVPRTVMIGGKAAPGYHMAKLIIKLVTSIGDVVNHDPV
      RatL     TMYNRIKK---------DPKKFFVPRTVIIGGKAAPGYHMAKMIIKLVTSVAEVVNNDPM
      HumL     TMYNRIKK---------DPKKLFVPRTVIIGGKAAPGYHMAKMIIKLITSVADVVNNDPM
      DictI    YRYLSIKKM-----SPKDRAQV-VPRVVIFAGKAAPGYVMAKRHIKLINSVAEVINRDKE
      DictII   NRYLDIK----------EGKKV-APRVVIFGGKAAPGYYMAKLIIKLINSVADVVNNDPK
      Yeast    YRYLAMKNMLKNGASIEEVAKKYPRKVSIFGGKSAPGYYMAKLIIKLINCVADIVNNDES
      PotH     YRYKKLKGM-----SPEERKNT-TPRTVMIAGKAFATYTNAKRIVKLVTDVGDVVNSDPD
      PotL     YRYKKMKEM-----TAAERKTNFVPRVCIFGGKAFATYVQAKRIVKFITDVGATINHDPE  <--
      EcglgP   TRYNRIKA---------DPDAKWVPRVNIFGGKAASAYYMAKHIIHLINDVAKVINNDPQ
      EcmalP   ALYKEIR---------ENPQADRVPRVFLFGAKAAPGYYLAKNIIFAINKVADVINNDPL
Conservation     i  ee          e       e  eeeeie e i  ii  e  e  ee  ei i
                    ff                  f  ffff f          f fff ff  f
      MamAnc   T YNRIKK---------DP KFFVPRTVIIGGKAAPGYHMAKMIIKLITSVADVVNNDP
      OvAnc    T Y RIK -------- DP  K VPRV IFGGKAAPGYYMAK II LI  VADV NNDP

Secondary      ....^^^^^^^...@@@@@@@@@@..^^^^^^^^......@@@@@@@@..^^^^^^^^.@
Buried         *   * * **     ***** ********************  *****************
Dimer
Binding                    vvv   v                 gggggvv  v
                    640       650       660       670       680       690
      RabM     VGDRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      RatM     VGDRFRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      HumM     VGDRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      RatB     VGDRLRVIFLENYRVSLAEKVIPAADLSQQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      HumB     VGDRLKVIFLENYRVSLAEKVIPAADLSQQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      RatL     VGSKLKVIFLENYRVSLAEKVIPATDLSEQVSTAGTEASGTGNMKFMLNGALTIGTMDGA
      HumL     VGSKLKVIFLENYRVSLAEKVIPATDLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      DictI    VDQYLKVVFIANYNVSIAQVIVPASDINQQISTAGTEASGTSNMKFTMNGSLIIGTLDGA
      DictII   VGDLLKVVFIPNYCVSNAEIIIPASDISQHISTAGTEASGTSNMKFSMNGGLIIGTLDGA
      Yeast    IEHLLKVVFVADYNVSKAEIIIPASDLSEHISTAGTEASGTSNMKFVMNGGLIIGTVDGA
      PotH     VNDYLKVVFVPNYNVSVAEMLIPGSELSQHISTAGMEASGTSNMKFALNGCLIIGTLDGA
      PotL     IGDLLKVVFVPDYNVSVAELLIPASDLSEHISTAGMEASGTSNMKFAMNGCIQIGTLDGA
      EcglgP   IGDKLKVVFIPNYSVSLAQLIIPAADLSEQISLAGTEASGTSNMKFALNGALTIGTLDGA
      EcmalP   VGDKLKVVFLPDYCVSAAEKLIPAADISEQISTAGKEASGTGNMKLALNGALTVGTLDGA
Conservation   e   eeieie ei ii ie eeieeeeeeei ii iiiiieiiie eii e eiieiii
               f   ff f f          ff f fff  f              f f   f f f
      MamAnc   VGDKLKVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGA
      OvAnc    VGDKLKVVFLPNY VSLAE  IPAADLSEQISTAGTEASGT NMKFALNGALTIGTLDGA
```

Fig. 1.

```
Secondary      @@@@@@@@...^^^^^^^^@@@@@@@@@@...@@@@@@@.@@@@@@@@@@@@.......
Buried         *****  *   *******    *  *      *      *  *  ** **  * **
Dimer
Binding
                   700       710       720       730       740       750
       RabM    NVEMAEEAGEENFFIFGMRVEDVDRLDQRGYNAQEYYDRIPELRQIIEQLSSGFFSPKQP
       RatM    NVEMAEEAGEDNFFIFGMRVEDVERLDQRGYNAQEYYDRIPELRQIIEQLSSGFFSPKQP
       HumM    NVEMAEEAGEENFFIFGMRVEDVDKLDQRGYNAQEYYDRIPELRQVIEQLSSGFFSPKQP
       RatB    NVEMAEEAGEENLFIFGMRVEDVEALDQKGYNAQEFYERLPELRQAVDQISSGFFSPKDP
       HumB    NVEMAEEAGAENLFIFGLRVEDVEALDRKGYNAREYYDHLPELKQAVDQISSGFFSPKEP
       RatL    NVEMAEEAGEENLFIFGMRVDDVAALDKKGYEAKEYYEALPELKLVIDQIDNGFFSPNQP
       HumL    NVEMAEEAGEENLFIFGMSIDDVAALDKKGYEAKEYYEALPELKLVIDQIDNGFFSPKQP
       DictI   NVEIAEEVGQENMFIFGLRTSEVEAAREKMTNKE----VNIDPRLQEVFLNIELGTFGPP
       DictII  NIEIRDAIGHENMYIFGARSEEVNKVKKIIHDGK----FTPDTRWARVLTAIKEDTFGPH
       Yeast   NVEITREIGEDNVFLFGNLSENVEELRYNHQYHPQ--DLPSSLDSVLSYIESGQFSPENP
       PotH    NVEIREEIGEDNFFLFGATADEVPQLRKDRENGLFKPDPRFEEAKQFIR-SGAFGTYDYN
       PotL    NVEIREEVGEENFFLFGAQAHEIAGLRKERADGKFVPDERFEEVKEFVR-SGAFGSYNYD
       EcglgP  NVEMLDHVGADNIFIFGNTAEEVEELRRQGYKPREYYEKDEELHQVLTQIGSGVFSPEDP
       EcmalP  NVEIAEKVGEENIFIFGHTVKQVKAILAKGYDPVKWRKKDKVLDAVLKELESGKYSDGDK
Conservation   ieie    i ei eeii    ee                                      e
                f f   f  f f f       f  f
       MamAnc  NVEMAEEAGEENLFIFGMRVEDVEALDKKGYNAKEYYD LPELRLVIDQISSGFFSPKQP
       OvAnc   NVE AE  GEEN FIFG   EEVEALR KGY   EYY    EL  VL QI SG FSP  P

Secondary      ...@@@@@@@@  @@^^^^@@@@@@@@@@@@@@@@@@@@@@.@@ @@@@@@@@@@@@@@@..
Buried           *  ** **     * *******  ** **  **          **  ** ****** **
Dimer
Binding
                   760       770       780       790       800       810
       RabM    DLFKDIVNML--MHHDRFKVFADYEEYVKCQERVSALYKNPR-EWTRMVIRNIATSGKFS
       RatM    DLFKDIVNMV--MHHDRFKVFADYEEYIKCQDKVSELYKNPR-EWTRMVIRNIATSGKFS
       HumM    DLFKDIVNML--MHHDRFKVFADYEDYIKCQEKVSALYKNPR-EWTRMVIRNIATSGKFS
       RatB    DCFKDVVNML--MYHDRFKVFADYEAYIQCQAQVDHLYRNPK-DWTKKVIRNIACSGKFS
       HumB    DCFKDIVNML--MHHDRFKVFADYEAYMQCQAQVDQLYRNPK-EWTKKVIRNIACSGKFS
       RatL    DLFKDIINML--FYHDRFKVFADYEAYVKCQEKVSQLYMNQK-AWNTMVLRNIAASGKFS
       HumL    DLFKDIINML--FYHDRFKVFADYEAYVKCQDKVSQLYMNPK-AWNTMVLKNIAASGKFS
       DictI   DVFRPILDSL--IFSDFYLSIQDFPLYLDSQASVDELWKDQS-AWVKKSIINSASTYFFS
       DictII  EQFQDIINSV-SGGNDHYILSYDFGSYLDIQNSIDQDFKDRA-KWAKKSIMASVCCGKFS
       Yeast   NEFKPLVDSIKYHGDYYLVSDDFESYYLATHELVDQEFHNQRSEWLKKSVLSVANVGFFS
       PotH    PLLESLEGNSGYGRGDYFLVGHDFPSYMDAQARVDEAYKDRK-RWIKMSILSTSGSGKFS
       PotL    DLIGSLEGNEGFGRADYFLVGKDFPSYIECQEKVDEAYRDQK-RWTTMSILNTAGSYKFS
       EcglgP  GRYRDLVDSL-INFGDHYQVLADYRSYVDCQDKVDELYELQE-EWTAKAMLNIANMGYFS
       EcmalP  HAFDQMLHSIGKQGGDPYLVMADFAAYVEAQKQVDVLYRDQE-AWTRAAILNTARCGMFS
Conservation       e                      fe  e  e           i    e        ii
                   f                       f     f                f
       MamAnc  DLFKDIVNML--  HDRFKVFADYEAY KCQEKV QLYKN K- WTKMVIRNIACSGKFS
       OvAnc     F DLVDSL-  GGD· LV AD    Y DCQ KVD LY DQ - WT   ILN A  G FS

Secondary      @@@@@@@@@@@@@.^^^^^^^^^^.....
Buried         ** *   **  **   *
Dimer
Binding
                       820       830       840
       RabM    SDRTIAQYAREIWGVEPSRQRLPAPDEKIP--------------------------------
       RatM    SDRTIAQYAREIWGLEPSRQRLPAPDEKI----------------------------------
       HumM    SDRTIAQYAREIWGVEPSRQRLPAPDEAI----------------------------------
       RatB    SDRTITEYAREIWGVEPSDLQIPPPNLPKD---------------------------------
       HumB    SDRTITEYAREIWGVEPSDLQIPPPNIPRD---------------------------------
       RatL    SDRTIREYAKDIWNMEPSDLKISLSKESSNGVNANGK--------------------------
       HumL    SDRTIKEYAQNIWNVEPSDLKISLSNES-NKVNGN----------------------------
       DictI   SDRAMNEYAEQIWDIKPCEVETTLNRRY---------------------------------
       DictII  SDRTIKEYAQQIWGIEEWKRPGPVPVSNEEARSLLVPPPSGSPNDINAISIERLSPLTFV
       Yeast   SDRCIEEYSDTIWNVEPVT-----------------------------------------
       PotH    SDRTISQYAKEIWNIAECRVP--------------------------------------
       PotL    SDRTIHEYAKDIWNIEAVEIA--------------------------------------
       EcglgP  SDRTIKEYADHIWHIDPVRL--------------------------------------
       EcmalP  SDRSIRDYQARIWQAKR----------------------------------------
Conservation   iii e ei   ii
                   f        f
       MamAnc  SDRTIKEYA  IWNVEPSDL I   NE --------------------------------
       OvAnc   SDRTIKEYA  IW I PV L  ----------------------------------------
```

Fig. 1.

```
Secondary
Buried
Dimer
Binding

        RabM    ----------------------------------------
        RatM    ----------------------------------------
        HumM    ----------------------------------------
        RatB    ----------------------------------------
        HumB    ----------------------------------------
        RatL    ----------------------------------------
        HumL    ----------------------------------------
        DICTI   ----------------------------------------
        DICTII  KQTSASPLSVISGGDKTNNTLKPKQTTKGFNIGGQPGNPTN
        Yeast   ----------------------------------------
        PotH    ----------------------------------------
        PotL    ----------------------------------------
        EcglgP  ----------------------------------------
        EcmalP  ----------------------------------------
Conservation

        MamAnc  ----------------------------------------
        OvAnc   ----------------------------------------
```

Figure 1. Alignment of the complete amino acid sequences of 14 phosphorylases. Sequences are from rabbit, rat and human muscle isozymes, RabM, RatM and HumM, respectively; rat and human brain isozymes, RatB and HumB, respectively; rat and human liver isozymes, RatL and HumL, respectively; phosphorylases 1 and 2 from *Dictyostelium*, DictI and DictII, respectively; yeast phosphorylase, Yeast, potato type H and L isozymes, PotH and PotL, respectively; and glycogen and maltodextrin phosphorylases from *E. coli*, EcglgP and EcmalP, respectively. The secondary structure of the rabbit M isozyme is shown in the row labeled "Secondary" and was determined from the crystal structure of the phosphorylated enzyme complexed with glucose (Sprang *et al.*, 1988; S. R. Sprang, E. J. Goldsmith and R. J. Fletterick, unpublished results) by a computer programme that determines main-chain angles (E. Fauman, unpublished results). The symbols @, . and ' represent alpha-helix, turn and beta-strand, respectively. The positions of buried residues, *, were obtained from Newgard *et al.* (1989) and are shown in the row labeled buried. Residues that form close contacts to residues in the other subunit (dimer contact residues), d, are indicated in the row labeled Dimer. Amino acid residues whose side-chains participate in ligand binding are indicated in the row labelled Binding and are marked: p, Ser-14 phosphorylation site; h, glucose-6-P binding site; a, AMP binding site; v, pyridoxal phosphate cofactor binding site; g, active site (glucose binding); s, glycogen storage site; and c, purine nucleoside (caffeine) inhibitor site. Three additional residues, aa71, aa309 and aa310, take part in glucose-6-P binding but are not identified as such in the Fig. since they also bind AMP. The positions of Ser-14 phosphate and caffeine binding residues were obtained from Newgard *et al.* (1989). AMP binding residues from Barford *et al.* (1991) and Sprang *et al.* (1987, 1988, 1991), residues that bind to oligosaccharides in the major and minor glycogen storage sites from Goldsmith *et al.* (1989*a*) and Johnson *et al.* (1990) and glucose-6-P binding residues from Johnson *et al.* (1992), Lorek *et al.* (1984) and Sprang *et al.* (1988). Residues in close contact (2·2 to 5·0 angstroms) with either glucose, pyridoxal phosphate or residues in the other subunit were determined from the phosphorylated rabbit M isozyme structure by the computer programme LIGPROT (S. R. Sprang, unpublished results). The positions of identical residues, i, conserved residues, e, (based primarily on codon conservation, from the program, CLUSTAL V; Higgins & Sharp, 1989) and conserved residues, f, (based on amino acid side-chain polarity, Rath *et al.*, 1987; Hudson *et al.*, 1993) are indicated in the row labeled Conservation. We define conserved residues, based on side-chain polarity, as those that fall within but not between the following groups of amino acids: (Gly, Ala, Val, Leu, Ile, Phe, Trp, Pro, Met, Cys), (Ser, Thr, Tyr, Asn, Gln), (Lys, Arg, His) and (Asp, Glu). Sequences of the mammalian and overall ancestral enzymes are shown in the rows labeled MamAnc and OvAnc, respectively.

In all 14 enzymes, 38·4% of the residues are either identical (19·4%), or conserved (19%), based on side-chain polarity. These residues are equally distributed between the N-terminal and C-terminal domains of the rabbit M isozyme. In the N-terminal domain, 19·3% of residues are identical and 18·5% are conserved and in the C-terminal domain, 19·6% of residues are identical and 19·6% are conserved. The activation subdomain within the N-terminal domain is less conserved containing 6·7% identical and 16·3% conserved residues. The majority (80·8%) of all identical and conserved residues occur at the positions of buried residues in the rabbit M isozyme and 61·3% of all buried residues in the rabbit M isozyme are either identical (30·9%) or conserved (30·4%) in all of the other enzymes.

The location of identical residues and residues that are conserved in a more evolutionary sense (based, in part, on codon conservation from the program CLUSTAL V (Higgins & Sharp, 1989)) is shown in Figure 2 in relation to the three-dimensional structure of the phosphorylated rabbit M isozyme complexed with glucose (Sprang *et al.*, 1988; S. R. Sprang, E. J. Goldsmith & R. J. Fletterick, unpublished results). The majority of identical residues are situated in a 10 to 15 angstrom radius surrounding the active site cleft of the enzyme. These include residues involved in binding substrate and the cofactor, pyridoxal phosphate, as well as residues that make up the secondary structural elements in which they are contained. The majority of conserved residues form an overlying

Table 1

Conservation of the overall structure of phosphorylase in relation to the primary sequence and three-dimensional structure of the rabbit M isozyme

Degree of conservation	Conserved residues				
	Total		Buried		
	(#)	(%)†	(#)	(% of total conserved)	(% of total buried)‡
Identical	158	19·4	127	40·7	30·9
Conserved	154	19·0	125	40·1	30·4
Total	312	38·4	252	80·8	61·3

The number (#) and percent (%) of identical and conserved (based on side-chain polarity) residues in all 14 phosphorylases is indicated.

† The total number of residues analyzed was 813.

‡ There are 411 buried residues in the rabbit muscle isozyme.

layer, approximately ten angstroms in thickness, which surrounds the identical residues at the active site cleft. Non-conserved residues and deletions/insertions present in the primary sequences of the non-mammalian enzymes are located primarily on the surface of the enzyme. Hence, the internal structure of the enzyme has been highly conserved.

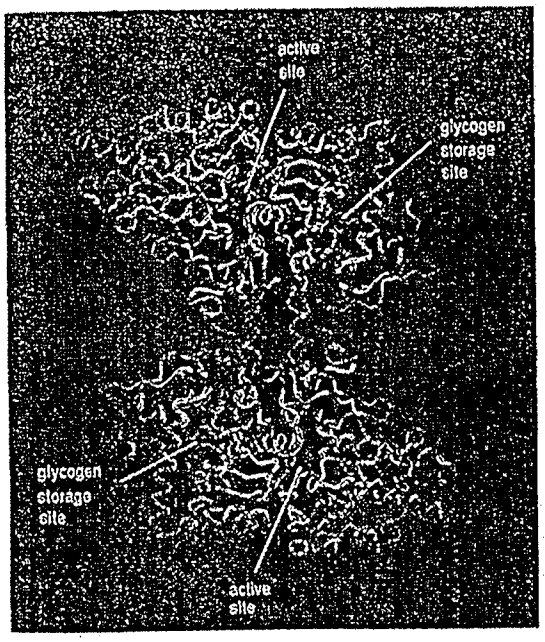


Figure 2. The location of identical (yellow), conserved (blue) and non-conserved residues (purple) throughout all 14 phosphorylases in relation to the 3-dimensional structure of the phosphorylated rabbit M isozyme complexed with glucose (Sprang *et al.*, 1988; S. R. Sprang, E. J. Goldsmith & R. J. Fletterick, unpublished results). Conserved residues were obtained using the program, CLUSTAL V (Higgins & Sharp, 1989). Residues were coloured using the program Insight II. The Fig. shows a ribbon model of the catalytic face of the dimer. The N and C-terminal domains are situated on the left and right-hand sides of the active site cleft, respectively.

Plots of the rates of change of the protein from the N-terminal to C-terminal ends in mammalian and non-mammalian phosphorylases are presented in Figure 3(a) and 3(b), respectively. The plots

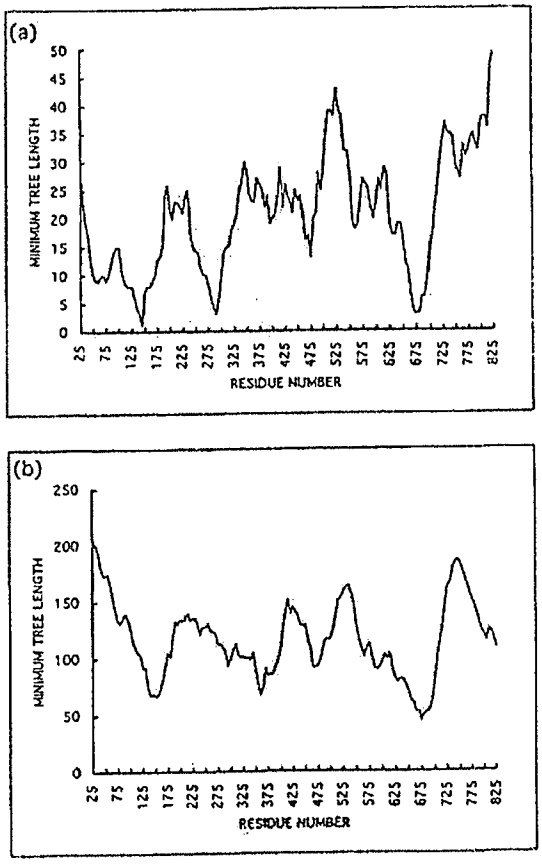


Figure 3. The rates of change of residues within phosphorylases. Plots represent the minimum tree length as a function of the residue number (rabbit M isozyme numbering). (a) Mammalian phosphorylases. (b) Non-mammalian phosphorylases.

represent estimates of minimum tree length as a function of residue number (G. B. Golding, unpublished results). Windows of 50 residues were examined at five residue intervals along the protein. Residues of phosphorylases that were included in the analysis correspond to those that align with the rabbit M isozyme sequence. Insertions in non-mammalian enzymes relative to the rabbit M isozyme sequence were not included in this analysis as their presence would lead to a false shortening of the tree length. Parts of the protein that are evolving at a slower rate than adjacent regions are indicated by troughs in the curves and those with lowest minimum tree length contain larger surrounding regions of conserved residues.

As expected, the minimum tree length for mammalian phosphorylases is much smaller than that for non-mammalian phosphorylases. This is due to the greater genetic distance between taxa. In the non-mammalian enzymes, the overall rate of change is relatively constant between the N-terminal and C-terminal domains. In contrast, the C-terminal domain of the mammalian isozymes appears to be changing at a greater rate overall than the N-terminal domain.

In mammalian phosphorylases, regions within the primary sequence that are evolving at a slower rate relative to adjacent areas occur around aa55 and aa75, which contain residues involved in AMP activation and glucose-6-P inhibition. Slower evolving regions also occur around aa145, aa285, aa385, aa480, aa565, aa600, aa635 and aa675, which contain residues within the active site, pyridoxal phosphate binding site and nucleoside inhibitor site of the rabbit M isozyme (see Fig. 1).

In non-mammalian enzymes, regions that are changing at a slower rate occur around aa45, aa75, aa290, aa360, aa385, aa465 to 475, aa565, aa585, aa630, aa675, aa805 and aa825. Most of these contain residues within the active site and pyridoxal phosphate binding site of the rabbit M isozyme. However, some also contain residues that are involved in allosteric control through either activation by AMP and/or inhibition by glucose-6-P (around aa45 and aa75). The slower evolving region around aa360 corresponds to a part of the minor glycogen storage site of the rabbit M isozyme. The slower rates of change around aa805 and aa825 may involve conservation of C-terminal residues that form the bottom part of the active site in the rabbit M isozyme, as suggested by Palm *et al.* (1987).

(b) *Phylogenetic relationships of phosphorylases*

A phylogeny based on the complete amino acid sequences of all 14 phosphorylases is presented in Figure 4. This phylogeny indicates that the two isozymes within either *E. coli*, potato or *Dictyostelium* are more closely related to each other than they are to phosphorylases from other species. Yeast phosphorylase is most closely related to the two phosphorylases from *Dictyostelium*. The long relative branch lengths of the non-mammalian

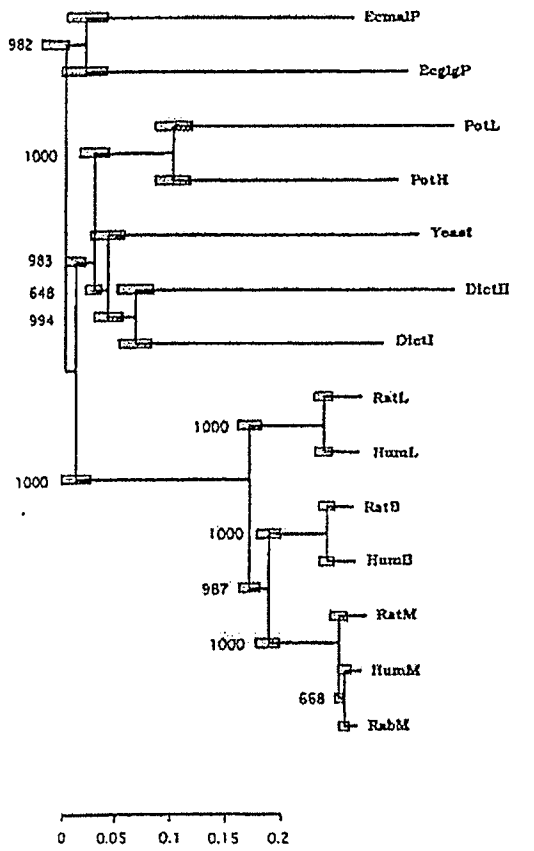


Figure 4. A phylogeny of phosphorylases based on their complete amino acid sequences. The numbers on the left indicate the number of bootstraps for which taxa to the right of that branch point form a monophyletic group. Branch lengths correspond to the median value and the stippled boxes represent 95% confidence intervals. The scale indicates the branch length for 0 to 0·2 amino acid replacements per site.

enzymes indicate that there has been either a high rate of change in these enzymes or that they are derived from ancient enzymes.

In mammals, the creation of the M, B and L isozymes predates the mammalian radiation. The M, B and L isozymes have evolved into phylogenetically distinct sequences and the M and B isozymes are more closely related to each other than to the L isozyme. The M isozyme is changing at a slower rate than the other two. This is a good indication that rates of change in tissue-specific isozymes can differ.

(c) *Conservation of ligand binding residues*

(i) *Active site and pyridoxal phosphate binding site*

As expected, active site and pyridoxal phosphate binding residues in the rabbit M isozyme are highly

conserved in both mammalian and non-mammalian phosphorylases (see Fig. 1). A total of 16 residues that come in close contact with glucose at the active site are identical in all phosphorylases with the exception of aa284 and aa484. Amino acid residue 284 is identical in all enzymes except the two potato isozymes where non-conservative changes occur (Asn/Glu and Asn/Ala) and in *E. coli* glycogen phosphorylase where there is a conservative change (Asn/Ser), based on side-change polarity. Amino acid residue 484 is identical in all enzymes except phosphorylase 2 from *Dictyostelium*, which has the substitution Asn→Ser. There is slightly more variation but still a high degree of conservation of the 15 residues that form close contacts with pyridoxal phosphate. Residues that are not identical are located at aa90, where all non-mammalian enzymes contain a non-conservative change (Tyr/Leu), at aa567, where there is a conservative change (Val/Ile) in a number of non-mammalian phosphorylase, at aa649, where non-conservative changes (Arg/Asn, Arg/Cys, Arg/Ser) occur in all non-mammalian phosphorylases and at aa677 where a non-conservative change (Gly/Ser) occurs in all non-mammalian enzymes except *E. coli* maltodextrin phosphorylase.

(ii) *Glycogen storage site*

Oligosaccharide binding to crystal structures of the rabbit M isozyme has indicated the existence of a minor (aa205 to 215 and aa354 to 362) and major (aa397 to 438) glycogen storage site in the protein (Goldsmith *et al.*, 1989*a*; Johnson *et al.*, 1990). A phylogeny based on all of the residues within the minor and major sites is presented in Figure 5. To allow a better comparison of corresponding residues in all phosphorylases, the 77 amino acid residue insert that is located in this region in the potato type L phosphorylase was omitted from the analysis. This phylogeny indicates that the glycogen storage site is continuing to change in the L isozyme from rat and human. In contrast, the glycogen storage site in the M and B isozymes is evolving at a much slower rate. As in the phylogeny of the overall sequences of these enzymes, the relative branch lengths of the non-mammalian phosphorylases are much greater than those of the mammalian isozymes for this region. Interestingly, apart from the 77 amino acid residue insertion, the type H and L isozymes from potato appear to be evolving at a similar rate although they differ dramatically in their affinities for glycogen (Mori *et al.*, 1991).

In mammalian phosphorylases, there is a high degree of conservation of those residues that correspond to sugar binding residues in the minor and major glycogen storage sites of the rabbit M isozyme (see Table 2). Human and rat M isozymes show 100% identity of all sugar binding residues, while B and L isozymes from human and rat show slightly less conservation of both sites (67 to 100% and 88 to 100% identical and conserved residues in the minor and major glycogen storage sites, respectively).

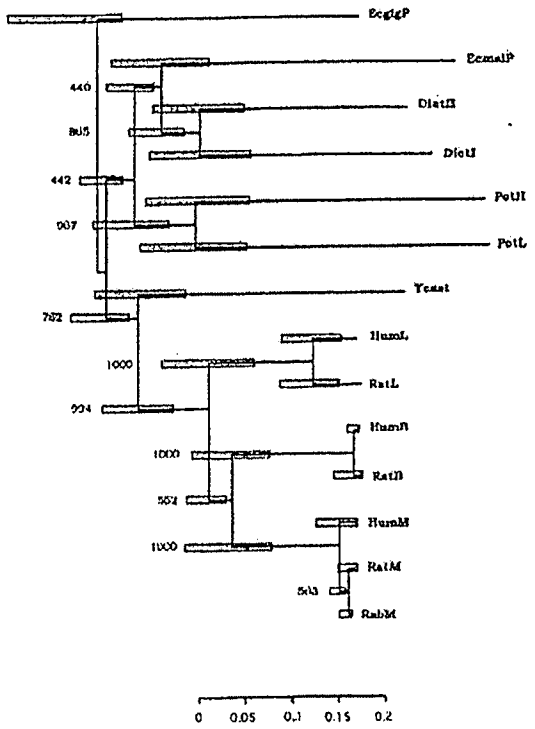


Figure 5. A phylogeny of phosphorylases based on residues corresponding to those in the major and minor glycogen storage sites of the rabbit M isozyme. The Fig. is drawn as described in Fig. 4.

In most cases, non-mammalian phosphorylases also show a relatively high degree of conservation of sugar binding residues in both sites (see Table 2). In the minor site, 50 to 83% of the sugar binding

Table 2

Conservation of sugar binding residues in the major and minor glycogen storage sites of the rabbit M isozyme

	Minor site†			Major site†		
Enzyme	I	C	%	I	C	%
Hum/rat M	6	0	100	17	0	100
Hum/rat B	4	2	100	14	1	88
Hum/rat L	3	1	67	14/13	2/4	94/100
Dict I	2	1	50	10	4	82
Dict II	2	2	67	11	2	76
Yeast	4	1	83	10	3	76
PotH	2	1	50	4	3	41
PotL	2	1	50	5	2	41
EcglgP	2	1	50	10	2	71
EcmalP	2	1	50	8	4	71

The number of identical, I, and conserved, C, (based on side-chain polarity), residues and the percent, %, of total conserved residues, I+C, is indicated.

† There are a total of 6 and 17 sugar binding residues in the minor and major glycogen storage sites, respectively.

residues are either identical or conserved, based on side-chain polarity. The C-terminal portion (aa354 to 362) of the minor site is much more conserved than the N-terminal portion (aa205 to 215). Of the five sugar binding residues in the C-terminal portion, 60 to 100% are either identical or conserved, depending on the enzyme (see Fig. 1). Three residues in the site, at aa354, aa355 and aa361, are either identical or conserved in all non-mammalian phosphorylases. In contrast, the sole sugar binding residue in the N-terminal portion of the minor site (aa213), is non-conserved in all non-mammalian phosphorylases. It is also non-conserved in the rat and human L isozymes. Deletions or insertions (1 to 6 aa in length) occur in this region in all non-mammalian enzymes except for the glycogen phosphorylase from *E. coli*. This indicates that the N-terminal portion of the minor site may not be important for polyglucan binding in many of these enzymes.

In the major site, highest conservation of sugar binding residues is observed in phosphorylase 1 and 2 from *Dictyostelium*, yeast phosphorylase and *E. coli* glycogen and maltodextrin phosphorylases. In these enzymes, 71 to 82% of the 17 sugar binding residues in the site are either identical or conserved. Sugar binding residues at aa398, aa402, aa407, aa431, aa433 and aa437 are identical or conserved in all five phosphorylases and sugar binding residues at aa405, aa411, aa425 and aa429 are identical or conserved in all but one of the five phosphorylases (see Fig. 1). Type H and L phosphorylases from potato exhibit a lower degree of conservation of sugar binding residues. In these enzymes, only 41% of the residues are either identical or conserved (see Table 2).

A long insertion (77 amino acid residues in length) occurs between aa420 and aa421 in the type L phosphorylase from potato (see Fig. 1). This insertion is situated in a loop on the surface of the protein that separates the two sugar binding helices of the major glycogen storage site in the rabbit M isozyme and it has been suggested that it may be responsible, in part, for the low affinity of the potato enzyme for glycogen (Nakano & Fukui, 1986). Small insertions (1 aa in length) between the sugar binding residues, aa433 and 437, exist in yeast phosphorylase and the two isozymes from potato. A deletion (3 aa in length) also occurs in this region in *E. coli* maltodextrin phosphorylase.

(iii) *Purine nucleoside inhibitor site*

The purine nucleoside inhibitor site of the rabbit M isozyme exists only in the inactive "T" conformational state of the enzyme and consists primarily of a hydrophobic slot constructed from Phe285 and Tyr613 (Kasvinsky *et al.*, 1978; Sprang *et al.*, 1982). This site is identically conserved in all mammalian phosphorylases, (see Fig. 1). In contrast, all non-mammalian phosphorylases, except the yeast enzyme, have non-conservative substitutions at aa285, indicating that this site may be non-functional in these enzymes. Tyr613 is identical in all phosphorylases. It has been suggested that this residue may also function to bind glycogen at the active site (Newgard *et al.*, 1989).

(iv) *Phosphorylation and AMP binding sites*

Ser14-P and AMP binding residues in the rabbit M isozyme are highly conserved in mammalian phosphorylases but poorly conserved in non-mammalian phosphorylases (see Table 3). The phosphate group of Ser14-P forms ion pairs with Arg43′ from the other subunit and Arg69 from the same subunit (Sprang *et al.*, 1988). (The prime symbol denotes residues in the other diad-related subunit.) These residues are identical in all of the mammalian isozymes. In contrast, non-conservative changes occur at the positions of either one or both residues in the non-mammalian enzymes.

There are 11 AMP binding residues in the rabbit M isozyme. These are situated in the "cap" loop of one subunit and helix-2, helix-8 and the aa315 to 325 loop of the other subunit (Barford *et al.*, 1991;

Table 3

Conservation of Ser14-P and AMP binding residues in the rabbit M isozyme

	Amino acid residue												
	Ser14-P		AMP										
Enzyme	43′	69	42′	44′	45′	71	72	75	309	310	315	316	318
Hum/rat M	I	I	I	I	I	I	I	I	I	I	I	I	I
Hum/rat B	I	I	I	I	I	I	I	I	I	I	I	I	I
Hum/rat L	I	I	I	I	I	I	I	I	I	I	I	I	NC
Diet I	C	NC	NC	C	NC	I	I	C	I	I	—	—	—
Yeast	NC	C	NC	C	NC	I	I	C	I	I	I	—	—
Diet II	C	NC	NC	C	NC	NC	NC	NC	NC	NC	—	—	—
PotH	NC	NC	NC	NC	NC	C	NC	NC	NC	I	NC	—	—
PotL	NC	NC	NC	NC	NC	C	NC	NC	NC	I	NC	—	—
EcglgP	NC	I	I	NC	I	C	NC	NC	NC	I	—	—	—
EcmalP	NC	NC	NC	NC	NC	NC	NC	NC	I	I	—	—	—

The symbols, I, C, NC and —, represent identical, conserved (based on side-chain polarity), non-conserved and deleted residues, respectively.

Sprang *et al.*, 1987, 1988, 1991). All of these residues are identical in the mammalian M and B isozymes (see Table 3). In the mammalian L isozyme, which is activated to a lesser extent by AMP, all AMP binding residues are identical, except aa318, which is not conserved (Cys/Ser) in both human and rat L isozymes.

Of the non-mammalian phosphorylases, only the *b* form of phosphorylase 1 from *Dictyostelium* is appreciably activated (6 to 8-fold) by AMP *in vitro* (Naranan *et al.*, 1988; Rutherford & Cloutier, 1986). In this enzyme, however, only four AMP binding residues are identical to those in the rabbit M isozyme. Two of these, Arg309 and Arg310, form ionic bonds in the rabbit M isozyme with the phosphate group of AMP, while the other two, Gln71 and Gln72, interact with the ribose group (Sprang *et al.*, 1987, 1988). Two residues are also conserved, based on side-chain polarity, representing Asn44 to Tyr44 and Tyr75 to Thr75 substitutions. In the latter substitution, the phenolic ring of Tyr75, which forms an important stacking interaction with the adenine base of AMP in the rabbit M isozyme, is missing in Thr75. This indicates that this stacking interaction is not preserved. Interestingly, yeast phosphorylase exhibits a very similar degree of conservation of AMP binding residues as the type 1 enzyme from *Dictyostelium*, but is not activated by AMP (Rath, 1991).

The remaining non-mammalian phosphorylases exhibit lower conservation of the AMP binding site, containing only one to four identical or conserved residues. Interestingly, all of the non-mammalian enzymes contain deletions in the aa315 to 325 loop of the rabbit M isozyme. This loop has been identified recently as an important binding determinant for the adenine base of AMP (Sprang *et al.*, 1991).

(v) *Glucose-6-P binding site*

In contrast to the Ser14-P and AMP binding sites, the glucose-6-P binding site is highly conserved in a number of non-mammalian phsophorylases (see Table 4). This is surprising since the glucose-6-P binding site overlaps with the AMP binding site in the rabbit M isozyme (Johnson *et al.*, 1992; Lorek *et al.*, 1984; Soprang *et al.*, 1988). Eight residues interact with glucose-6-P and three of these are also involved in AMP binding. Two of these, Arg309 and Arg310 interact with the common phosphate groups of AMP and glucose-6-P. The other, Gln71, interacts with the sugar groups of both ligands. The remaining five residues in the glucose-6-P binding site form further contacts with the sugar (Val40′, Trp67, Arg193, Asp227) and phosphate (Arg242) groups.

Table 4
Conservation of glucose-6-P binding residues in the rabbit M isozyme

	Amino acid residue							
Enzyme	40′	67	71	193	227	242	309	310
Hum/rat M	I	I	I	I	I	I	I	I
Hum/rat B	I	I	I	I	I	I	I	I
Hum/rat L	I	I	I	I	I	I	I	I
DictL	C	I	I	I	I	I	I	I
YeastL	C	I	I	I	I	I	I	I
PotL	I	I	C	I	I	I	NC	I
EcglgP	C	I	C	I	I	I	NC	I
DictII	C	I	NC	I	I	I	NC	NC
PotH	NC	C	C	I	I	I	NC	I
EcmalP	C	C	NC	NC	I	I	I	I

Symbols are as described in Table 3.

All glucose-6-P binding residues are identical in the rat and human M, B and L isozymes, indicating a high degree of conservation of the site in mammalian phosphorylases (see Table 4). There is also strong conservation of the site in four non-mammalian phosphorylases. Seven of the eight glucose-6-P binding residues in the rabbit M isozyme are identical in the *Dictyostelium* type 1 and yeast phosphorylases. The other residue, at aa40′, is conserved, being a Val to Ala change in both enzymes. The type L phosphorylase from potato and the glycogen phosphorylase from *E. coli* also exhibit a high degree of conservation of the site. Six and five glucose-6-P binding residues are identical in the potato and *E. coli* enzymes, respectively, and only one residue is non-conserved, being a Arg/Ser change at aa309 in both enzymes. This non-conservative substitution may have less effect on glucose-6-P binding than on AMP binding since the phosphate group of glucose-6-P not only forms ion pairs with Arg309 and Arg310, but also interacts with Arg242 (Johnson *et al.*, 1992; Lorek *et al.*, 1984; Sprang *et al.*, 1988). The other three non-mammalian enzymes exhibit lower conservation, containing two to three non-conservative changes in residues that bind both the phosphate and sugar groups of the ligand. Interestingly, an intrasubunit hydrogen bond network containing residues Tyr157, Asp227, Arg242, Asp306 and Arg310 is formed at the phosphate subsite (Johnson *et al.*, 1992) and this network is also highly conserved. All of the residues in the network are identical in all of the non-mammalian enzymes with the exception of *Dictyostelium* type 2 phosphorylase, which contains a non-conservative substitution at aa310.

(d) *Conservation of dimer contact residues*

To obtain estimates of the degree of conservation of dimer contact residues in the rabbit M isozyme, these residues were first separated into interacting groups based on the residues that they contact in the other subunit. This analysis indicated that they could be grouped into three relatively independent networks (see Fig. 6). Interestingly, one network appears to be associated with phosphorylation/AMP control as it contains dimer contact and other closely linked residues that bind Ser14-P and AMP (aa42′ to 45′, aa69, aa71, aa72 and aa75). The second network contains dimer contact residues involved in glucose-6-P binding (aa40′, aa67 and aa193) and hence, may be more closely associated

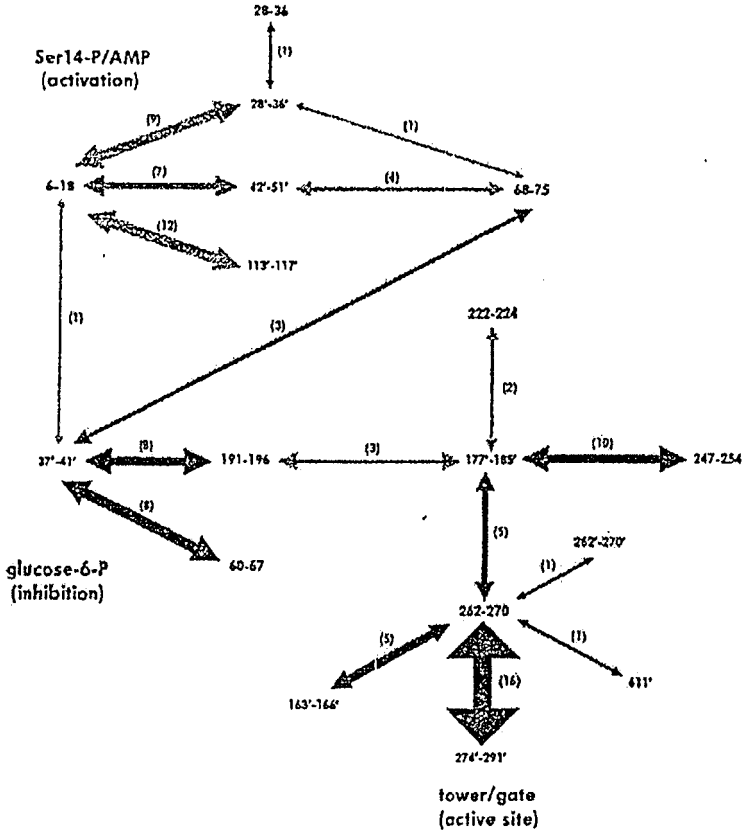


Figure 6. Conservation of dimer contact networks in the rabbit M isozyme. Interacting groups of dimer contact residues are separated by arrows. The number of dimer contact pairs in each interacting group is given in brackets. The degree of conservation of dimer contact pairs with interacting groups is indicated by the degree of shading of the arrows: light grey, 0 to 15%, medium grey, 16 to 35% and dark grey, 36 to 60%.

with glucose-6-P control. The third network contains dimer contact residues within the tower helix (aa262 to 267) and on both sides of the gate loop (aa282' to 285') in the other subunit. These elements are directly associated with conformational changes at the active site (Barford & Johnson, 1989; Johnson *et al.*, 1992; Sprang *et al.*, 1987). Dimer contact pairs within each interacting group are presented in Table 5.

The degree of conservation of these dimer contact networks was determined by examining the conservation of interacting pairs of dimer contact residues. Our criteria for conservation of an interacting pair is that both dimer contact residues of the pair be identical or conserved (based on side-chain polarity) relative to the corresponding residue in the rabbit M isozyme. All mammalian isozymes exhibit a high degree of conservation of dimer contact pairs (see Fig. 1 and Hudson *et al.*, 1993).

Table 6 presents estimates of the degree of conservation of these dimer contact networks among the non-mammalian phosphorylases. The network associated with Ser14-P and AMP binding residues is poorly conserved. On average, only three dimer contact pairs are conserved out of a total of 34 in the network. Type 1 phosphorylase from *Dictyostelium*, which is activated by both phosphorylation and AMP, and yeast phosphorylase, which is activated by phosphorylation but not AMP, fall within the range of conservation exhibited by the other enzymes. Four dimer contact pairs that form a link (link 1) between the networks associated with Ser14-P/AMP and glucose-6-P binding residues are also poorly conserved.

In contrast, dimer contact pairs in the network associated with glucose-6-P binding residues exhibit a much higher degree of conservation. Phosphorylase 1 from *Dictyostelium*, yeast phosphorylase, type L phosphorylase from potato and glycogen phosphorylase from *E. coli* exhibit highest conservation, containing 9 or 10 conserved pairs out of 16 in the network. Phosphorylase 2 from *Dictyostelium*, type H phosphorylase from potato and *E. coli* maltodextrin phosphorylase have less conserved networks, containing two to six conserved dimer contact pairs. For this non-mammalian enzyme set, the relative degree of conservation of the network correlates with the relative

Table 5

Dimer contact pairs within interacting groups in the rabbit M isozyme

Network	Amino acid residues	
	Interacting groups	Dimer contact pairs†
Ser14-P/AMP	6–18/28′–36′	12/28′, 12/32′, 13/32′, 13/35′, 13/36′, 14/32′, 14/36′, 15/36′, 16/32′
	6–18/42′–51′	10/43′, 10/51′, 11/43′, 12/43′, 13/43′, 13/51′, 14/43′
	6–18/113′–117′	6/116′, 9/113′, 9/114′, 9/115′, 9/116′, 10/115′, 10/116′, 10/117′, 12/114′, 12/115′, 13/115′, 13/117′
	28′–36′/28–36	33′/33
	28′–36′/68–75	36′/68
	42′–51′/68–75	42′/68, 42′/72, 44′/72, 45′/75
Link 1‡	6–18/37′–41′	18/37′
	37′–41′/68–75	40′/68, 40′/71, 41′/68
Glucose-6-P	37′–41′/60–67	37′/60, 37′/61, 37′/64, 37′/65, 38′/60, 40′/60, 40′/64, 40′/67
	37′–41′/191–196	38′/191, 39′/191, 39′/193, 40′/191, 40′/193, 41′/193, 41′/195, 41′/196
Link 2‡	177′–185′/191–196	184′/194, 185′/194, 185′/195
Tower/gate	177′–185′/222–224	184′/222, 184′/224,
	177′–185′/247–254	177′/254, 178′/250, 178′/251, 178′/254, 179′/250, 179′/254, 181′/247, 181′/250, 184′/247, 184′/250
	177′–185′/262–270	177′/262, 178′/262, 178′/266, 179′/269, 181′/269
	163′–166′/262–270	163′/266, 163′/269, 164′/262, 165′/262, 166′/262
	262–270/274′–291′	262/278′, 262/279′, 262/281′, 263/278′, 263/279′, 263/280′, 263/281′, 263/291′, 266/277′, 266/278′, 267/274′, 267/277′, 267/278′, 267/291′, 270/277′, 270/274′
	262–270/262′–270′	270/270′
	262–270/611′	262/611′

† Dimer contact pairs (2·2 to 5·0 angstroms apart) were determined from the crystal structure of the phosphorylated rabbit M isozyme complexed with glucose (Sprang *et al.*, 1988; S. R. Sprang, E. J. Goldsmith and R. J. Fletterick, unpublished results) using the computer programme LIGPROT (S. R. Sprang, unpublished results).

‡ Link 1 and link 2 represent interacting pairs between the Ser14-P/AMP and glucose-6-P associated networks and between the glucose-6-P and tower helix/gate loop associated networks, respectively.

Table 6

Conservation of dimer contact pairs of the rabbit M isozyme in non-mammalian phosphorylases

Network†	Enzyme	Conserved dimer contact pairs‡	
		(#)	(%)
Ser14-P/AMP (34)	Diet I	6	18
	Yeast	4	12
	All others	3·0 (1–6)	9 (3–18)
Link 1 (4)	All	0·9 (0–3)	23 (0–75)
Glucose-6-P (16)	Diet I	9	56
	Yeast	9	56
	PotL	10	63
	EcglgP	10	63
	DictII	6	38
	PotH	2	13
	EcmalP	3	31
Link 2 (3)	All	0·4 (0–2)	13 (0–67)
Tower/gate 1 (17)	All	4·3 (2–8)	25 (12–47)
Tower/gate 2 (23)	PotL	1	4
	EcamlP	5	22
	All others	13·8 (10–16)	60 (43–70)

† Interacting groups of dimer contact residues within networks are as described in Table 5. The total number of dimer contact pairs within a network is indicated in brackets.

‡ When more than 1 phosphorylase is considered, the number (#), and percent (%), of conserved dimer contact pairs represents an average with the range given in brackets.

degree of conservation of the glucose-6-P binding site.

Based on the degree of conservation of interacting groups, the tower helix and gate loop associated network was divided into two parts. Part 1 of the network (containing interacting groups: aa177′ to 185′, aa222 to 224, aa247 to 254 and aa262 to 270) is poorly conserved. In this part of the network, the dimer contact groups, with the exception of aa177′ to 185′, do not interact directly with the groups (aa262 to 270 and aa274′ to 291′) that contain tower helix and gate loop residues (see Fig. 6). Part 2 of the network (containing interacting groups: aa163′ to 166′, aa262 to 270, aa274′ to 291′ and aa611′) exhibits greater conservation in all non-mammalian phosphorylases except the type L phosphorylase from potato and *E. coli* maltodextrin phosphorylase. On average, 13·8 dimer contact pairs out of 23 are conserved in the other five non-mammalian enzymes.

The degree of conservation of single, interacting groups in all seven non-mammalian phosphorylases is illustrated in Figure 6. Interacting groups with the highest degree of conservation (36 to 60% conserved pairs) are those associated with glucose-6-P binding residues (aa37′ to 41′/aa191 to 196 and aa37′ to 41′/aa60 to 67) and those directly associated with tower helix and gate loop residues (aa262 to 270/aa163′ to 166′, aa262 to 270/aa274′ to 291′, aa262 to 270/aa262′ to 270′ and aa262 to 270/

aa611'). Interacting groups with a lower degree of conservation (16 to 35% conserved pairs) occur within the Ser14-P/AMP associated network (aa28 to 36/aa28' to 36', which contains only one dimer contact pair, aa33/aa33'), one of the links between the Ser14-P/AMP and glucose-6-P associated networks (aa37' to 41'/aa68 to 75) and within part 1 of the tower helix/gate loop associated network (aa177' to 185'/aa262 to 270 and aa177' to 185'/aa247 to 254). All other interacting groups exhibit poor conservation (0 to 15% conserved pairs).

(c) *Conservation of ligand binding sites and dimer contact networks in the mammalian and overall ancestral sequences*

Reconstructed sequences of the mammalian and overall ancestral enzymes are presented in Figure 1. The overall ancestral sequence was derived from the sequences of all the mammalian and non-mammalian ancestral enzymes in the phylogeny presented in Figure 4. The mammalian ancestral sequence was derived from the mammalian set of ancestral enzymes. The extent of conservation of residues comprising ligand binding sites and dimer contact networks is shown in Table 7.

As expected, the mammalian and overall ancestral enzymes contain a high degree of conservation of active site and pyridoxal-phosphate binding residues. Residues within the major and minor glycogen storage sites are also relatively highly conserved in both ancestral sequences. All caffeine-binding residues in the purine nucleoside inhibitor site are identical in the mammalian ancestral enzyme. However, in the overall ancestral sequence, a non-conservative substitution (Phe/Thr) exists at aa285 (see Fig. 1) indicating that the nucleoside inhibitor site might not have been functional in the overall ancestral enzyme.

In the mammalian ancestor, there is high conservation of residues in the Ser14-P and AMP binding sites. In contrast, the overall ancestral enzyme exhibits poor conservation of these sites. For the Ser14-P binding site, only one of the two residues that interact with the phosphate group is conserved (Arg69). For the AMP binding site, only three of 11 AMP binding residues are identical (Val45', Arg309, Arg310). The remaining residues are either unknown or not present in the overall ancestral sequence. There is also no conservation in the overall ancestral enzyme of dimer contact pairs in the phosphorylation/AMP associated dimer contact network. In contrast, this network is highly conserved in the mammalian ancestral enzyme.

Both mammalian and overall ancestral enzymes exhibit high conservation of the glucose-6-P binding site. All eight glucose-6-P binding residues are identical in the mammalian ancestral sequence and six of these are indentical in the overall ancestral sequence. These six residues bind to both the phosphate (Arg242, Arg309 and Arg310) and sugar groups (Trp67, Arg193 and Asp227) of glucose-6-P. There is also some conservation in the overall ancestral sequence of interacting pairs of residues in the glucose-6-P associated dimer contact network. Five out of the 16 dimer contact pairs are identical in the network and four of these (out of 8) are present within the interacting group, aa37' to 41'/aa60 to 67.

Dimer contact pairs are also conserved in the tower helix/gate loop associated network. In the overall ancestral sequence, 41% and 70% of dimer contact pairs are identical or conserved in parts 1 and 2 of the network, respectively. Highly conserved interacting groups in the network are aa177' to 185'/aa247 to 254, aa177' to 185'/aa262 to 270, aa163' to 166'/aa262 to 270, aa262 to 270/aa274' to 291', aa262 to 270/aa262' to 270' and aa262 to 270/aa611'. In these groups, 4/10 3/5, 2/5, 12/16, 1/1, 1/1 dimer contact pairs are conserved, respectively.

Table 7

The number (#) and percent (%) of identical and conserved (based on side-chain polarity) ligand binding residues and dimer contact pairs in the mammalian and overall ancestral sequences

Binding site/ dimer contact network	Mammalian ancestor		Overall ancestor	
	(#)	(%)	(#)	(%)
A. *Ligand binding sites*				
Active site (glucose)	16	100	16	100
Pyridoxal-phosphate	15	100	11	73
Glycogen storage site				
(major)	15	88	12	71
(minor)	5	83	3	50
Purine nucleoside inhibitor site	4	100	3	75
Ser14-P	2	100	1	50
AMP	10	91	3	27
Glucose-6-P	8	100	6	75
B. *Dimer contact networks*†				
Ser14-P/AMP	33	97	0	0
Link 1	3	75	0	0
Glucose-6-P	16	100	5	31
Link 2	3	100	0	0
Tower/gate (part 1)	16	94	7	41
Tower gate (part 2)	21	91	16	70

† Interacting groups in dimer contact networks are as presented in Table 5.

4. Discussion

The alignment presented in Figure 1 is in agreement at the N and C termini with that obtained by Newgard *et al.* (1989) for a subset of the mammalian isozymes, the two phosphorylases from *E. coli*, the yeast enzyme and the type L enzyme from potato. It differs, however, from alignments obtained for the type H enzyme from potato (Mori *et al.*, 1991) and the type 1 and 2 enzymes from *Dictyostelium* (Rogers *et al.*, 1992; Rutherford *et al.*, 1992). Our

alignment also differs from that of Newgard *et al.* (1989) and Nakano & Fukui (1986) with regard to the position and length of the large insert in the type L enzyme from potato. There are also minor differences in the positions of several insertions/deletions in the yeast enzyme in comparison to the alignment of Newgard *et al.* (1989). The greater number of phosphorylase sequences used in this study provides a more extensive data set for matching amino acid residues. This has led to the construction of an alignment that provides a higher degree of sequence similarity among these enzymes than observed previously in certain regions of the protein. These include regions near the N terminus, as well as the portion of the protein that corresponds to the major glycogen storage site in the rabbit M isozyme.

As expected, the internal core of the protein, as well as active site and pyridoxal phosphate binding residues, are highly conserved. The N-terminal domain exhibits a similar degree of overall conservation as the C-terminal domain when all 14 sequences are compared. In contrast, previous comparisons with subsets of these sequences have indicated, in general, that the C-terminal domain is more conserved than the N-terminal domain (Newgard *et al.*, 1989). Our analysis of the relative overall rates of change of residues in non-mammalian organisms also indicates that both domains are evolving at a similar rate. This suggests that the selective pressures on both domains are equivalent. Interestingly, the C-terminal domain is evolving at a faster rate in mammalian phosphorylases than the N-terminal domain. This may reflect the greater use of allosteric control in these enzymes.

Analysis of the degree of conservation of residues corresponding to oligosaccharide binding residues in the major and minor glycogen storage sites of the rabbit M isozyme (Goldsmith *et al.*, 1989*a*; Johnson *et al.*, 1990) indicate that both sites are well conserved in the mammalian and most of the non-mammalian enzymes. The type L and H phosphorylases from potato exhibit the lowest degree of conservation. Both enzymes contain 50% and 41% conserved sugar binding residues in the minor and major sites, respectively. Since the two enzymes contain a similar degree of conservation of these two sites, the lower affinity of the type L enzyme for glycogen is likely due in large part, to the large 77 residue insertion in the major glycogen storage site, as suggested previously (Nakano & Fukui, 1986).

Surprisingly, sugar binding residues in the major glycogen storage site are more conserved (71% conservation) in *E. coli* maltodextrin phosphorylase than they are in the high glycogen affinity, type H enzyme from potato. Maltodextrin phosphorylase also exhibits a similar degree of conservation of residues in both the major and minor sites as the true glycogen phosphorylase from *E. coli*. Interestingly, a three amino acid deletion in maltodextrin phosphorylase occurs between two important sugar binding residues, aa433 and aa437. This region in the rabbit M isozyme undergoes a major conformational change upon oligossacharide binding (Goldsmith *et al.*, 1989*a*; Johnson *et al.*, 1990) and has been implicated in oligosaccharide length specificity (Johnson *et al.*, 1990). Hence, it is possible that this deletion in maltodextrin phosphorylase may be responsible in part for the enzyme's preference for maltodextrin as substrate.

The portion of the enzyme that is shared by all of these phosphorylases appears to span nearly the entire length (95%) of the rabbit M isozyme with the exception of the N-terminal tail, which contains residues in phosphorylation control, and the C terminus, which contains residues that are displaced by the N terminus upon phosphorylation (Barford *et al.*, 1991; Sprang *et al.*, 1991). Sequence identity and conservation of amino acid residues begins at helix-1 (aa23 to 40) and extends at least to the end of the last alpha-helix, helix 33 (aa813 to 825) in relation to the rabbit M isozyme (see Fig. 1). In helix-1, all three buried residues (aa27, aa31 and aa35) are conserved, containing non-polar side-chains in all of the enzymes. In addition, all of the enzymes, except glycogen and maltodextrin phosphorylase from *E. coli*, contain a His residue at aa34 and all of the enzymes, except *E. coli* maltodextrin phosphorylase, contain a Thr residue at aa38. Glycogen and maltodextrin phosphorylases from *E. coli* contain conserved Lys and Tyr substitutions at aa34 and aa38, respectively. In addition to these primary sequence considerations, helix-1 has been observed in the three-dimensional structure of yeast phosphorylase *b* (V. L. Rath & R. J. Fletterick unpublished results). The appearance of identical residues in all 14 enzymes begins at aa63 in helix-2. In helix-33, there are six residues that are identical in all enzymes. Five of these occur at residues that are buried in the rabbit M isozyme and two, aa824 and aa825, are located at the C-terminal end of the helix. The sequence of the overall ancestral enzyme derived from all mammalian and non-mammalian ancestral enzymes also begins in helix-1, at aa25, and extends to aa832. This ancestral enzyme is similar, but not identical, in sequence to the mammalian ancestral enzyme at both the N and C termini, as well as in internal regions throughout the protein.

It has been postulated that the N-terminal 80 amino acid residues in the mammalian phosphorylases may have been acquired by fusion of a unique segment on to a mammalian ancestral gene (Newgard *et al.*, 1986, 1989). Our results do not support this hypothesis, but indicate that allosteric control involving ligand binding and dimer contact residues within helix-1, the CAP loop of the AMP activation site (aa41′ to aa47′) and helix-2 (aa48 to 78), evolved from a single ancestral gene common to all of the enzymes examined here. This does not rule out the possibility of gene fusion events occurring before helix-1, since this could explain the different sites and mechanisms for phosphorylation of the yeast and mammalian enzymes, as has been suggested previously (Hwang & Fletterick, 1986).

In the rabbit M isozyme, there are 74 dimer

contact residues that interact with residues in the other subunit. As these residues are important for the propagation of allosteric effects (Barford *et al.*, 1991; Sprang *et al.*, 1988, 1991), we have carried out a detailed analysis of the degree of conservation of dimer contact pairs in all of these phosphorylases. To make greater sense of the reason underlying conservation of specific dimer contact pairs, dimer contact residues in the rabbit M isozyme were first divided into interacting groups based on the residues that they contact in the other subunit. Surprisingly, this analysis indicated that interacting groups could be placed into three relatively independent networks: a network containing phosphorylation and AMP binding residues, a network containing glucose-6-P binding residues and an active site associated network, containing residues within the tower helix and gate loop. Since relatively few dimer contact pairs exist between these networks, this raises the possibility that enzyme activation by phosphorylation and AMP might involve the use of an intersubunit pathway for allosteric signal transmission that is distinct from that used for inhibition of enzyme activity by glucose-6-P.

In support of this hypothesis, the degree of conservation in non-mammalian phosphorylases of dimer contact pairs in the networks associated with either phosphorylation/AMP or glucose-6-P binding residues correlates with the degree of conservation of the sites to which these ligands bind. Phosphorylation and AMP binding site residues are poorly conserved in the non-mammalian enzymes and poor conservation is also observed for the associated dimer contact network. Phosphorylase 1 from *Dictyostelium* is the only enzyme in the non-mammalian set examined here that is significantly activated by AMP (Naranan *et al.*, 1988; Rutherford & Cloutier, 1986). As the AMP binding site and associated dimer contact network are poorly conserved, a separate pathway for allosteric signal transmission across subunits may have evolved in this enzyme through convergent evolution. A different pathway to that used in mammals might also have evolved for phosphorylation control in both the *Dictyostelium* 1 and yeast enzymes.

In contrast, glucose-6-P binding residues are highly conserved in four of the non-mammalian enzymes. All residues comprising the glucose-6-P binding site are either identical or conserved in phosphorylase 1 from *Dictyostelium*, yeast phosphorylase, type L phosphorylase from potato and *E. coli* glycogen phosphorylase, with the exception of a non-conservative substitution that occurs at aa309 in the latter two enzymes. The glucose-6-P binding site is less well conserved in phosphorylase 2 from *Dictyostelium*, type H phosphorylase from potato and *E. coli* maltodextrin phosphorylase. In these two groups of non-mammalian enzymes, the relative degree of conservation of dimer contact pairs in the network associated with glucose-6-P binding residues correlates with the relative degree of conservation of the glucose-6-P binding residues. This provides evidence in favor of a possible functional role for this network in glucose-6-P control.

The high degree of conservation of the glucose-6-P binding site and associated dimer contact network in the non-mammalian enzymes favors a model for the evolution of allostery of the enzyme where control by glucose-6-P existed earlier than control by either phosphorylation or AMP. In accord with this, the glucose-6-P binding site and associated dimer contact network is more conserved in the overall ancestral sequence than the phosphorylation/AMP binding sites and associated dimer contact network.

It is possible that the AMP binding site in mammals evolved initially through the use of a portion of the phosphate binding subsite for glucose-6-P involving residues Arg309, Arg310, as suggested previously (Newgard *et al.*, 1989). However, residues chosen for binding the sugar and base moities of AMP were distinct, for the most part, from those used for binding the sugar moiety of glucose-6-P. This perhaps allowed the establishment of a distinct dimer contact network that could then be used as a separate pathway for the propagation of allosteric signals leading to activation rather than inhibition of enzyme activity. The relatively high degree of conservation of the dimer contact network associated with the tower helix and gate loop in the non-mammalian enzymes and in the overall ancestral enzyme suggests that these structural elements may have been involved early on in promoting conformational changes at the active site. Confirmation of this model will require detailed kinetic analysis of the effect of glucose-6-P on the activity of a variety of non-mammalian enzymes. Crystallographic studies will also be necessary for identifying residues involved in dimer formation and ligand binding. In this respect, projects are underway for determining the three-dimensional structures of phosphorylases from potato and yeast (Hecht *et al.*, 1987; Rath *et al.*, 1992). The functional independence of glucose-6-P and phosphorylation/AMP associated dimer contact networks could be assessed by examining the effects of mutations leading to the creation or destruction of contacts between specific dimer contact pairs in the rabbit M isozyme.

We gratefully acknowledge Drs R. J. Beynon, M. F. Browner, R. J. Fletterick, P. K. Hwang and V. L. Rath for helpful advice throughout the preparation of this manuscript and E. C. Gardonio for help with the preparation of the Figures. This work was supported by NSERC grants, A8383 and U0336 to M.M.C. and G.B.G., respectively. Support was also provided by a NATO Collaborative Research Grant to M.M.C. and by the Canadian Institute for Advanced Research to G.B.G.

References

Archarya, K. R., Stuart, D. I., Varvill, K. M. & Johnson, L. N. (1991). *Glycogen Phosphorylase b. Description of the Protein Structure*, World Scientific Publishing Co. Pte. Ltd., Singapore.

Barford, D. & Johnson, L. N. (1989). The allosteric transition of glycogen phosphorylase. *Nature (London)*, 340, 609–616.

Barford, D., Hu, S.-H. & Johnson, L. N. (1991). Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP. *J. Mol. Biol.* 218, 233–260.

Browner, M. F. & Fletterick, R. J. (1992). Phosphorylase: a biological transducer. *Trends Biochem. Sci.* 17, 66–71.

Browner, M. F., Fauman, E. B. & Fletterick, R. J. (1992*a*). Tracking conformational states in allosteric transitions of phosphorylase. *Biochemistry*, 31, 11297–11304.

Browner, M. F., Hwang, P. K. & Fletterick, R. J. (1992*b*). Cooperative binding is not required for activation of muscle phosphorylase. *Biochemistry*, 31, 11291–11296.

Burke, J., Hwang, P., Anderson, L., Lebo, R., Gorin, F. & Fletterick, R. (1987). Intron/exon structure of the human gene for the muscle isozyme of glycogen phosphorylase. *Proteins: Struct. Funct. Genet.* 2, 177–187.

Choi, Y.-L., Kawamukai, M., Utsumi, R., Sakai, H. & Komano, T. (1989). Molecular cloning and sequencing of the glycogen phosphorylase from *Escherichia coli*. *FEBS Letters*, 243, 193–198.

Coats, W. S., Browner, M. F., Fletterick, R. J. & Newgard, C. B. (1991). An engineered liver glycogen phosphorylase with AMP allosteric activation. *J. Biol. Chem.* 266, 16113–16119.

Crerar, M. M., Hudson, J. W., Matthews, K. E., David, E. S. & Golding, G. B. (1988). Studies on the expression and evolution of the glycogen phosphorylase gene family in the rat. *Genome*, 30, 582–590.

Felsenstein, J. (1990). PHYLIP manual. Version 3.3 Herbarium, University of California, Berkeley.

Fukui, T., Shimomura, S. & Nakano, K. (1982). Potato and rabbit muscle phosphorylases: comparative studies on the structure, function and regulation of regulatory and nonregulatory enzymes. *Mol. Cell. Biochem.* 42, 129–144.

Gelinas, R. P., Froman, B. E., McElroy, F., Tait, R. C. & Gorin, F. A. (1989). Human brain glycogen phosphorylase: characterization of fetal cDNA and genomic sequences. *Mol. Brain Res.* 6, 177–185.

Glaser, T., Matthews, K. E., Hudson, J. W., Seth, P., Housman, D. E. & Crerar, M. M. (1989). Localization of the muscle, liver and brain glycogen phosphorylase genes on linkage maps of mouse chromosomes 19, 12 and 2, respectively. *Genomics*, 5, 510–521.

Goldsmith, E. J., Sprang, S. R. & Fletterick, R. J. (1989*a*). Alternative binding modes for maltopentaose in the activation site of glycogen phosphorylase. *Trans. Amer. Chem. Assoc.* 25, 87–104.

Goldsmith, E. J., Sprang, S. R., Hamlin, R., Xuong, N.-H. & Fletterick, R. J. (1989*b*). Domain separation in the activation of glycogen phosphorylase *a*. Science, 245, 528–532.

Hecht, H. J., Jahnke, K. & Buehner, M. (1987). Crystallization and structure determination of potato phosphorylase at 3 Angstrom resolution. In *Biochemistry of Vitamin B6* (Korpela, T. & Christen, P. eds), pp. 287. Birkhauser Verlag, Basel.

Higgins, D. G. & Sharp, P. M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS*, 5, 151–153.

Hudson, J. W., Hefferon, K. L. & Crerar, M. M. (1993). Comparative analysis of species-independent, isozyme-specific amino acid substitutions in mammalian muscle, brain and liver glycogen phosphorylases. *Biochim. Biophys. Acta*, 1164, 197–208.

Hwang, P. K. & Fletterick, R. J. (1986). Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases. *Nature (London)*, 324, 80–84.

Johnson, L. N., Acharya, K. R., Jordan, M. D. & McLaughlin, P. J. (1990). Refined crystal structure of the phosphorylase-heptulose 2-phosphate-oligosaccharide-AMP complex. *J. Mol. Biol.* 211, 645–661.

Johnson, L. N., Barford, D., Acharya, R., Oikonomakos, N. K. & Martin, J. L. (1992). Allosteric regulation of glycogen phosphorylase. In *Proceedings of the Robert A. Welch Foundation Conference on Chemical Research XXXVI. Regulation of Proteins by Ligands*, pp. 17–35. Robert A. Welch Foundation, Houston, Texas.

Kasvinsky, P. J., Madsen, N. B., Sygusch, J. & Fletterick, R. J. (1978). The regulation of glycogen phosphorylase *a* by nucleotide derivatives: kinetic and X-ray crystallographic studies. *J. Biol. Chem.* 253, 3343–3351.

Lorek, A., Wilson, K. S., Sansom, M. S. P., Stuart, D. I., Stura, E. A., Jenkins, J. A., Zanotti, G., Hajdu, J. & Johnson, L. N. (1984). Allosteric interaction of glycogen phosphorylase *b*. A crystallographic study of glucose-6-P and inorganic phosphate binding to di-imidate-cross-linked phosphorylase *b*. *Biochem. J.* 218, 45–60.

Mori, H., Tanizawa, K. & Fukui, T. (1991). Potato tuber type H phosphorylase isozyme. Molecular cloning, nucleotide sequence, and expression of a full-length cDNA in *Escherichia coli*. *J. Biol. Chem.* 266, 18466–18453.

Nakano, K. & Fukui, T. (1986). The complete amino acid sequence of potato alpha-glucan phosphorylase. *J. Biol. Chem.* 261, 8230–8236.

Nakkano, K., Hwang, P. K. & Fletterick, R. J. (1986). Complete cDNA sequence for rabbit muscle glycogen phosphorylase. *FEBS Letters*. 204, 283–287.

Nakano, K., Mori, H. & Fukui, T. (1989). Molecular cloning of cDNA encoding potato amyloplast alpha-glucan phosphorylase and the structure of its transit peptide. *J. Biochem.* 106, 691–695.

Naranan, V., Brickey, D. A. & Rutherford, C. L. (1988). Glycogen phosphorylase "b" in *Dictyostelium:* stability and endogenous phosphorylation. *Mol. Cell. Biochem.* 83, 89–104.

Newgard, C. B., Nakano, K., Hwang, P. K. & Fletterick, R. J. (1986). Sequence analysis of the cDNA encoding human liver glycogen phosphorylase reveals tissue-specific codon usage. *Proc. Nat. Acad. Sci., U.S.A.*, 83, 8132–8136.

Newgard, C. B., Littman, D. R., van Cenderen, C., Smith, M. & Fletterick, R. J. (1988). Human brain glycogen phosphorylase: cloning, sequence analysis, chromosomal mapping, tissue expression and comparison with human liver and muscle isozymes. *J. Biol. Chem.* 263, 3850–3857.

Newgard, C. B., Hwang, P. K. & Fletterick, R. J. (1989). The family of glycogen phosphorylases: structure and function. *Crit. Rev. Biochem. Mol. Biol.* 24, 69–99.

Palm, D., Goerl, R. & Burger, K. J. (1985). Evolution of catalytic and regulatory sites in phosphorylases. *Nature (London)*, 313, 500–502.

Palm, D., Goerl, G., Weidinger, G., Zeier, B., Fischer, B. & Schinzel, R. (1987). *E. coli* maltodextrin phosphorylase: primary structure and deletion mapping of the C-terminal site. *Z. naturforsch*, 42c, 394–400.

Rath, V. L. (1991). Allosteric mechanisms in phosphorylases. Ph.D. thesis, University of California, San Francisco.

Rath, V. L., Newgard, C. B., Sprang, S. R., Goldsmith, E. J. & Fletterick, R. J. (1987). Modeling the biochemical differences between rabbit muscle and human liver phosphorylase. *Proteins: Struct. Funct. Genet.* 2, 225–235.

Rath, V. L., Hwang, P. K. & Fletterick, R. J. (1992). Purification and crystallization of glycogen phosphorylase from *Saccharomyces cerevisiae*. *J. Mol. Biol.* 225, 1027–1034.

Rogers, P. V., Luo, S., Sucic, J. F. & Rutherford, C. L. (1992). Characterization and cloning of glycogen phosphorylase 1 from *Dictyostelium discoideum*. *Biochim. Biophys. Acta*, 1129, 262–272.

Rutherford, C. L. & Cloutier, M. J. (1986). Identification of two forms of glycogen phosphorylase in *Dictyostelium*. *Arch. Biochem. Biophys.* 250, 435–439.

Rutherford, C. L., Peery, R. B., Sucic, J. F., Yin, Y., Rogers, P. V., Luo, S. & Selmin, O. (1992). Cloning, structural analysis and expression of the glycogen phosphorylase-2 gene in *Dictyostelium*. *J. Biol. Chem.* 267, 2294–2302.

Sagardia, F., Gotay, J. & Rodriquez, M. (1971). Control properties of yeast glycogen phosphorylase. *Biochem. Biophys. Res. Commun.* 42, 829–835.

Saitou, N. & Nei, M. (1987). The neighbour joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Schiebel, K., Pekel, E. & Mayer, D. (1992). The nucleotide sequence of rat liver glycogen phosphorylase cDNA. *Biochim. Biophys. Acta*, 1130, 349–351.

Shimomura, S. & Fukui, T. (1980). A comparative study on alpha-glucan phosphorylases from plant and animal: interrelationship between the polysaccharide and pryidoxal phosphate binding sites by affinity electrophoresis. *Biochemistry*, 19, 2287–2294.

Shimomura, S., Nagai, M. & Fukui, T. (1982). Comparative glucan specificities of two types of spinach leaf phosphorylase. *J. Biochem.* 91, 703–717.

Sprang, S., Fletterick, R., Stern, M., Yang, D., Madsen, N. & Sturtevant, J. (1982). Analysis of an allosteric binding site: the nucleoside inhibitor site of phosphorylase *a*. *Biochemistry*, 21, 2036–2048.

Sprang, S. R., Goldsmith, E. & Fletterick, R. J. (1987). Structure of the nucleotide activation switch in glycogen phosphorylase *a*. *Science*, 237, 1012–1019.

Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., Madsen, N. B. & Johnson, L. N. (1988). Structural changes in glycogen phosphorylase induced by phosphorylation. *Nature (London)*, 336, 215–221.

Sprang, S. R., Withers, S. G., Goldsmith, E. J., Fletterick, R. J. & Madsen, N. B. (1991). Structural basis for the activation of glycogen phosphorylase *b* by adenosine monophosphate. *Science*, 254, 1367–1371.

Yu, F., Jen, Y., Takeuchi, E., Inouye, M., Nakayama, H., Tagaya, M. & Fukui, T. (1988). Alpha-glucan phosphorylase from *Escherichia coli*, cloning of the gene, and purification and characterization of the protein. *J. Biol. Chem.* 263, 13706–13711.

Edited by A. Klug